بسم الله الرحمن الرحيم

جامـــــعـــة اليرمـــــــوك

كـــليـــــــة التربـــيـــــة

قسم علم النفس الإرشادي والتربوي



التحقق من فاعلية مؤشرات الكشف عن افتراض أحادية البعد وفق نماذج نظرية استجابة الفقرة، في ضوء تغير طول الاختبار وشكل توزيع القدرة.

Verifying the Effectiveness of Unidimensionality Assumption Indicators According to Item Response Theory Models (IRT), in light of test length and Ability distribution.

إعداد:

أروى عيسى عبد الحواري

إشراف الدكتور:

نضال كمال شريفين

حقل التخصص: القياس والتقويم

2013/2012

قرار لجنة المناقشة

التحقق من فاعلية مؤشرات الكشف عن افتراض أحادية البعد وفق نماذج نظرية استجابة الفقرة، في ضوء تغير طول الاختبار وشكل توزيع القدرة.

إعداد:

أروى عيسى عبد الحواري

بكالوريوس فيزياء، جامعة اليرموك، 1992

دبلوم تربية، أساليب تدريس العلوم، جامعة اليرموك، 2002

ماجستير قياس وتقويم، جامعة اليرموك، 2008

قدمت هذه الأطروحة استكمالا لمتطلبات الحصول على درجة دكتوراه الفلسفة في تخصص القياس والتقويم في جامعة اليرموك، إربد، الأردن.

وافق عليها

د. نضال كمال شريفين .... رئيساً

أستاذ مشارك في القياس والتقويم التربوي، جامعة اليرموك.

أ.د. أحمد سليمان عودة .... عضواً

أستاذ البحث والتقويم التربوي، جامعة جدارا.

أ.د. أحمد يوسف قواسمة .... عضواً

أستاذ القياس والتقويم التربوي، جامعة اليرموك.

أ.د. شفيق فلاح علاونة .... عضواً

أستاذ علم النفس التربوي، جامعة اليرموك.

أ.د. غازي ضيف الله رواقة .... عضواً

أستاذ مناهج وطرق تدريس التعلم المهني، جامعة اليرموك.

الإهداء

إلى منْ نقش الحرف في حياتي: والدي (عيسى الحواري) رحمه الله

إلى تلك التي لا تزال تضيء دربي: أمي حفظها ربي

إلى أسرتي وقرة عيني (زوجي) و(أبنائي) حفظهم ربي

إلى كل من دعمني ووقف إلى جواري

إلى كل طالب علم

أهدي هذا الجهد المتواضع

الباحثة

شكر وتقدير

الحمد لله الذي أعانني على إتمام هذا العمل، ومنَ علي من فضله الكريم لتخرج هذه الرسالة إلى حيز الوجود، والصلاة والسلام على سيدنا محمد خير الأنام عليه أفضل الصلاة والتسليم، لا يسعني إلا أن أتقدم بجزيل الشكر والعرفان وعظيم الامتنان إلى أستاذي الفاضل الكريم، الدكتور نضال كمال شريفين، الذي أعطاني الكثير من علمه الوافر، وفتح لي أبواب العلم والمعرفة ولم يتوانى لحظة عن مد يد العون والمساعدة، وتقديم الدعم والمتابعة بكـل صدر رحب على الرغم من أعبائه الكبيرة، فكل حرف من حروفهـا يـدين لـه بالفـضل والعرفان، فجزاه الله عني كل خير وحفظه ورعاه.

كما أتقدم بالشكر للأساتذة الأفاضل أعضاء لجنة المناقشة، الأستاذ الـدكتور أحمـد عودة، والأستاذ الدكتور أحمد قواسمة، والأستاذ الدكتور شفيق علاونة، والأستاذ الـدكتور غازي رواقه، والذين تفضلوا مشكورين بقبول مناقشة هذه الرسالة أثابهم الله عني كل خير.

و أتقدم بالشكر للدكتور محمود القرعان لما قدمه لي من عون ومساعدة، جزاه الله عني كل خير. وجزيل شكري وعظيم امتناني لأسرتي وأبنائي لما تحمّلوه معي من تعـب وعناء حفظهم الله ورعاهم، وتقديري الكبير إلى إخواني وأخواتي الأعزاء، وإلى كل مـن تعاون معي في إخراج هذه الرسالة، جزاهم الله عني كل خير.

فهرس المحتويات

فهرس الجداول

فهرس الأشكال

الشكل الصفحة

الملخص

الحواري، أروى عيسى. التحقق من فاعلية مؤشرات الكشف عن افتراض أحادية البعد وفق نماذج نظرية استجابة الفقرة، في ضوء تغير طول الاختبار وشكل توزيع القدرة. رسالة دكتوراه، جامعة اليرموك 2013 (المشرف: د. نضال كمال شريفين).

هدفت الدراسة الحالية للتحقق من فاعلية مؤشرات الكشف عن افتراض أحادية البعد وفق نماذج نظرية استجابة الفقرة، في ضوء تغير طول الاختبار وشكل توزيع القدرة. باستخدام التحليل العاملي التوكيدي. ولتحقيق هذا الهدف تم تعريف افتراض أحادية البعد وبيان أهميته، والآثار الناجمة عن انتهاكه. وتم التعريف أيضا بمؤشرات الكشف عن أحادية البعد، كما تم تعريف نماذج استجابة الفقرة والافتراضات التي تقوم عليها، إضافة إلى التعريف بنموذج المعادلة البنائية وبيان أهميته، والشروط اللازمة لإجراء التحليل العاملي التوكيدي كأحد تطبيقاته، وبيان أهم المؤشرات التي تنتج عنه وهي مؤشرات جودة المطابقة، إضافة إلى التعريف بأهم خصائص تلك المؤشرات.

وتمت مناقشة كافة مؤشرات جودة المطابقة وهي: النسبة بين قيمة χ^2 ودرجات الحرية df ومؤشرات المطابقة المطلقة (Absolute Fit Index) ، ومؤشرات المطابقة المتزايدة Incremental Fit Indexes ، إضافة إلى مؤشر هولتر لجودة المطابقة بمستوى Hoelter's (1983)`critical N' for a 0.05 . وفقاً لعلامة قطع متفق عليها في الأدب النظري.

كشفت نتائج التحليل العاملي التوكيدي عن أهمية افتراض أحادية البعد، وعن فاعلية المؤشرات المستخدمة في الكشف عن افتراض أحادية البعد. ومن ضمن إجراءات هذه الدراسة تم توليد البيانات، وتم توليد مجموعة الفقرات، كما تم توليد قدرات للأفراد، ولتحقيق ذلك تم توليد بيانات بواقع (1000) مفحوص لكافة متغيرات الدراسة؛ حيث طول الاختبار (15، 30، 60) فقرة، وشكل توزيعات القدرة (طبيعية، ملتو التواءً موجباً وملتو التواءً سالباً)، وفق نموذج

استجابة الفقرة (أحادي المعلمة، وثنائي المعلمة، وثلاثي المعلمة)، وتم تحليل البيانات المولدة باستخدام برمجية AMOS 21. كما استخدم الاختبار مربع كاي(χ^2) لحسن المطابقة للكشف عن فاعلية مؤشرات جودة المطابقة.

أظهرت النتائج أن لطول الاختبار تأثيراً في فاعلية المؤشرات المستخدمة في التحقق من افتراض أحادية البعد، وأن طول الاختبار 30 فقرة حقق تطابقا بين النموذج المقترح والبيانات المولدة؛ وأظهرت المؤشرات الإحصائية AGFI، GFI، RMR، فاعليتها في الكشف عن افتراض أحادية البعد، كما كان لشكل توزع القدرة أثر في فاعلية تلك المؤشرات، حيث وجد أن شكل توزيع القدرة الملتوي التواءً سالباً والملتوي التواءً موجباً يحققان افتراض أحادية البعد، وأن للنماذج اللوجستية أثراً في فاعلية المؤشرات المستخدمة، إذ أشارت النتائج إلى فاعلية المؤشر RMR في حالة البيانات المولدة من النموذج اللوجستي ثلاثي المعلم.

وفي ضوء النتائج توصي الباحثة باستخدام المؤشرات الآتية (RMR، GFI، AGFI) وهي مؤشرات جودة المطابقة التابعة للتحليل العاملي التوكيدي في الكشف عن افتراض أحادية البعد كأسلوب جديد في الكشف عن افتراض أحادية البعد، وذلك بغض النظر عن العوامل المؤثرة في افتراض أحادية البعد، كونها تصلح كمؤشرات للكشف عن أحادية البعد، علاوة على ذلك فهي ذات أقل تأثر بحجم الاختبار وشكل توزيع القدرة ونوع النموذج المستخدم.

الكلمات المفتاحية: افتراض أحادية البعد، تحليل عاملي توكيدي، نموذج المعادلة البنائية، مؤشرات جودة المطابقة، نموذج استجابة الفقرة، برمجية AMOS.

الفصل الأول

خلفية الدراسة وأهميتها

المقدمة:

نظرا لأهمية الاختبارات وبنوك الأسئلة في مجال قياس السلوك الإنساني والسلوك النفسي وما ينطوي عليها من قرارات ونتائج وتعديل وتطوير؛ تعددت مجالات البحث المتعلقة بها وبكيفية بنائها وكيفية تطويرها. كما تعددت استخداماتها في المجالات النفسية والتربوية، وأصبح يعتمد على العديد منها في اتخاذ القرارات الضرورية في كافة القطاعات العامة والخاصة، وأهمها قطاع وزارة التربية والتعليم؛ إذ يُعتمد عليها في العديد من التطبيقات التربوية كاختبارات القدرة المتنوعة واختبارات الكفاءة وغيرها من الاختبارات المختلفة؛ المرجو تحقيق الموضوعية منها في القياس المطلوب.

ولا تزال الاختبارات تتمتع بأهميتها في مجالات التقييم في مدارس وزارة التربية والتعليم، ولا يزال البحث منصباً عليها لأهميتها في القياسات النفسية والتربوية، وخاصة تلك التي يتم بناؤها في ضوء نماذج رياضية مختلفة ترتبط بالنظرية الحديثة في القياس (IRT) التي يعوّل عليها كثيرا في بناء الاختبارات وبنوك الأسئلة الحديثة، بحيث تتمتع بخصائص سيكومترية توفر لمستخدمي ومطوري الاختبارات وصانعي القرارات الخيارات المتعددة في اختيار فقرات المقياس المحدد والمطلوب حسب غايته وهدفه.

إلا أنه يوجد بعض المتغيرات التي يمكن أن تؤثر في دقة الفقرات وصدق قياسها، بسبب تجاوز بعض الافتراضات الأساسية الخاصة بالنظرية الحديثة في القياس والتي يجب أن تتمتع بها البيانات الخاصة بالفقرات التي تستخدم في بناء مثل هذه البنوك. ومن هذه الافتراضات: افتراض أحادية البعد والاستقلال الموضعي والسرعة، التي من شأنها تهديد صدق البناء

للمقاييس المختلفة؛ حيث أشارت العديد من الدراسات السابقة (Daniel, Allen, & Jams, 2002 ;Douglas, 2006 ;Wang, 2005; Reese & Pashley, 2007) التي أجريت في هذا النطاق إلى تأثير انتهاك هذه الافتراضات على الخصائص السيكومترية لتلك الاختبارات والتي يجرى عليها التحليلات باستخدام أحد النماذج الرياضية المرتبطة بنظرية استجابة الفقرة.

وتعتبر هذه الافتراضات – أحادية البعد، والاستقلال الموضعي، والسرعة – غاية في الأهمية، ويعد انتهاكها تجاوزا خطيرا في مجالات القياس النفسي والتربوي، الذي ينعكس سلباً على القرارات الضرورية التي من شأنها الارتقاء بمجالات العلوم التربوية والمقاييس النفسية. ولافتراض أحادية البعد أهمية لا تقل شأنا عن الافتراضات الأخرى، في الكشف عن السمة المقاسة، فقد ذكر لي وتيري (Lee & Terry, 2005) أن انتهاك افتراض أحادية البعد، يؤثر سلباً على الصدق البنائي للاختبار، الذي يعتمد على البناء النظري للسمات التي تقيسها فقرات الاختبار، فيكون انتهاك افتراض أحادية البعد المصدر الرئيس لانخفاض قيمته. فعندما يقيس الاختبار أكثر من سمتين كامنتين، يكون ترتيب المفحوصين على بعد واحد (سمة واحدة) غير متحقق، ويكون الكشف عن تلك السمة غير متحقق بسبب وجود أبعاد أخرى (سمات أخرى) في الفضاءات الكامنة المتعددة، وبالتالي ينعكس هذا على صدق البناء في نموذج استجابة الفقرة أحادية البعد.

وأشار ستاوت (Stout, 1987) إلى ضرورة أهمية أحادية البعد في التطبيقات الاختبارية المتعددة، للحفاظ على ترتيب الأفراد في المقياس الواحد، إضافة إلى تجنب التحيز الناشئ عن تقديرات معالم الفقرة وتقديرات قدرات الأفراد.

وبين ووكر وبرتراس (Walker & Beretras, 2000) أن تطبيق بيانات متعددة الأبعاد في نموذج يفترض أحادية البعد يؤدي إلى زيادة في خطأ القياس، وبالتالي يقود إلى نتائج غير

صحيحة حول معالم الفقرة، ومعلم القدرة للأفراد، ويعطي تقديرا متحيــزا للثبــات ودالــة المعلومات للاختبار، وان استخدام نموذج استجابة الفقرة أحادي البعد على بيانات متعــددة الأبعاد يعد استنتاجا غير صحيح لتوزيع قدرات المفحوصين على الأبعاد الثانوية.

وقد وجد أن معدي ومطوري الاختبارات يتجاهلون تلك الافتراضات أثناء بنائها مما يوقعهم في العديد من أخطاء القياس التي تنعكس سلبا على النتائج؛ والتي تبنى عليها العديد من القرارات، فيغيب عن الأذهان ضرورة التأكد من أبعاد الاختبار والسمات التي يجب أن يكشف عنها قبل بنائها. من هنا كان لا بد من التأكيد على افتراض أحادية البعد وأهميته إلى جانب الافتراضات الأخرى للتحقق من صدق البناء وموضوعية القياس(Hattie, 1985).

وهناك ضرورة للتحقق من افتراض أحادية البعد وكيفية التحقــق منــه، لأهميتــه الكبيرة في مجال القياس والبحوث المختلفة، على الرغم من أن افتــراض أحاديــة البعــد يصعب تحقيقه بشكل تام لأنه يفسر وجود قدرة واحدة فقط يمكن أن تفسر أداء الفــرد فــي اختبار ما، إذ أن هناك العديد من العوامل التي تؤثر في استجابة الأفراد كالدافعيــة والقلــق والمستوى المعرفي وغيرها. ويتطلب تحقيق هذا الفرض وجود عامل مسيطر واحد علــى الاختبار وهو القدرة المقاسة (علام، 2005).

ويرى هامبلتون وسواميناثان (Hambleton & Swaminathan, 1985) ضــرورة أن تقيس الفقرة بعدا واحدا على السمة المقاسة، وإذا تعرضت لأي خلل؛ فإن الفقرة تقيس بعداً آخراً غير الذي يقصده كاتب الفقرة، وهذا يؤثر في عملية صدق القياس، إضافة إلى عــدم قدرتها في الكشف عن السمة المراد قياسها بصورة حقيقية فينعكس ذلك على ثبات وصدق المقاييس. إن افتراض أحادية البعد شرطا أساسيا في بعض نماذج استجابة الفقرة. وهذا الشرط مطلوب في اختبارات القدرة.

وتتأثر أساسيات افتراض أحادية البعد بظروف وحجم الاختبار، فقــد بـــين زانـــج (Zhang, 2008)، أن الاختبارات القصيرة أكثر عرضة لانتهاك افتراض أحادية البعد، لـــذا عندما يتم انتهاك افتراض أحادية البعد فإن الاختبارات القصيرة تتأثر كثيرا في قيـــاس مـــا أعدت لقياسه.

وقد ذكر كارن ومايكل وكـــالن (Karon, Michael, & Kallen, 2008) أن طبيعـــة توزيع البيانات تؤثر على افتراض أحادية البعد، وأن حجم الفقرات الكبير يقـــاوم انتهـــاك التوزيع الطبيعي للبيانات، ونقصان حجم الفقرات في الاختبارات يتسبب في انتهـــاك هـــذا الافتراض، كما أن هذا الافتراض وثيق الصلة بتوازن أو تجانس المحتوى.

وقد بين هاملبتون وروفنيل (Hambleton & Rovinelli, 1986) وجود عـــدد مـــن الأساليب والطرق لإثبات افتراض أحادية البعد .وتتراوح هذه الأساليب من مقاييس الاتساق الداخلي المستخدمة بصورة شائعة إلى التمثيل البياني لقيم الجذور الكامنة (Eigenvalue Plot) والإحصاءات المرتبطة بها، حيث أحصى الباحثان السابقان عددا من المؤشرات في أدبيـــات القياس النفسي للتحقق من افتراض أحادية البعد الخاصة بفقرات الاختبار، ومع ذلك فإن العديد من هذه الأساليب لم تأخذ حقها في الدراسة كما يجب.

وقد أكد هاتي (Hattie, 1985) على ضـــرورة التحقـــق مـــن فاعليـــة المؤشـــرات المستخدمة في الكشف عن أحادية البعد في ظل الظروف المختلفـــة فـــي تـــصميم وبنـــاء الاختبارات وتطويرها؛ من اجل إيجاد الاتساق بين الأساليب المختلفة التـــي تـــستخدم فـــي الكشف عن أحادية البعد، لذا جاءت هذه الدراسة لتحقق هذا الغرض في الكشف عن فاعلية مؤشرات الكشف عن افتراض أحادية البعد تبعاً لنماذج نظرية استجابة الفقـــرة فـــي ظـــل ظروف مختلفة لتغير طول الاختبار وشكل توزيع القدرة.

مشكلة الدراسة وأسئلتها:

يعد افتراض أحادية البعد من الافتراضات الهامة التي يجب عدم تجاهلها في مجالات القياس والاختبارات المتعددة، وأن عدم الالتزام بها يوقع معدي الاختبارات ومطوري المقاييس في العديد من أخطاء القياس التي تنعكس سلبا على النتائج؛ فتقود إلى نتائج غير صحيحة حول معالم الفقرة، ومعلم القدرة للأفراد التي تبنى عليها العديد من القرارات الهامة، فيقع على كاهلهم ضرورة التأكد من أبعاد الاختبار والسمات التي يجب أن يكشف عنها قبل بنائها. من هنا كان لا بد من التأكيد على افتراض أحادية البعد وأهميته للتحقق من صدق البناء وموضوعية القياس.

كما وتأتي ضرورة إجراء الدراسة الحالية لبيان أهمية افتراض أحادية البعد على دقة تدريج فقرات الاختبارات، لتحقيق أقصى درجات الموضوعية في مجال القياس السلوكي والإنساني، وهذا يعود بالنفع على المؤسسات العلمية المختلفة التي تهتم باستخدام وإنشاء الاختبارات المختلفة وكذلك البحوث العلمية التي تتعلق بأغراض القياس المختلفة بالاعتماد على تلك الاختبارات. إضافة للتحقق من دقة المؤشرات المستخدمة في الكشف عن افتراض أحادية البعد ليصبح لدى مطورو الاختبارات مؤشرات يمكن بواسطتها التحقق من افتراض أحادية البعد، وأنه في حالة عدم تواجد مثل تلك المؤشرات سيهدد صدق البناء للمقاييس المختلفة، وهذا يتسبب في كثير من المشاكل لمعدي ومطوري الاختبارات وصناع القرار.

لذا جاءت هذه الدراسة للكشف عن تحقق افتراض أحادية البعد باستخدام التحليل العاملي التوكيدي الذي لم يأخذ حقه من الدراسة في المقاييس والبحوث المختلفة، وخاصة أنه لم يتم استخدام إلا عدد قليل من المؤشرات التي لم تكشف بدقة عن افتراض أحادية البعد، أتت

هذه الدراسة لتلقي الضوء على فاعلية المؤشرات المتبعة في الكشف عن افتراض أحادية البعد وفق نماذج استجابة الفقرة تبعا لتغير طول الاختبار وشكل توزيع القدرة، وتحديدا لتجيب عن الأسئلة التالية:

السؤال الأول:

هل تختلف فاعلية مؤشرات الكشف عن تحقق افتراض أحادية البعد في النموذج البنائي المستخدم باختلاف طول الاختبار (15، 30، 60) فقرة؟

السؤال الثاني:

هل تختلف فاعلية مؤشرات الكشف عن تحقق افتراض أحادية البعد في النموذج البنائي المستخدم باختلاف شكل توزيع القدرة (التوزيع الطبيعي، ملتوٍ التواءً موجباً، ملتوٍ التواءً سالباً)؟

السؤال الثالث:

هل تختلف فاعلية مؤشرات الكشف عن تحقق افتراض أحادية البعد في النموذج البنائي المستخدم باختلاف النموذج (أحادي المعلمة، ثنائي المعلمة، ثلاثي المعلمة)؟

أهمية الدراسة:

تأتي أهمية هذه الدراسة من أهمية افتراض أحادية البعد خاصة في مجال الاختبارات والمقاييس المختلفة كأحد الافتراضات الضرورية الواجب التأكد منها، لذا تسعى هذه الدراسة للتحقق من افتراض أحادية البعد الذي يؤكد على قياس بعد واحد وتفسير سمة واحدة تتعلق بمقياس لقدرة واحدة، وهذا يساعد على إيجاد المقياس الصحيح والذي من خلاله يمكن الكشف عن السمة بدقة وعناية، كما ويساعد تحقيق افتراض أحادية البعد على إمكانية بناء صور متعددة للمقاييس من خلال معايرتها ومعادلتها، لضمان صدق المقياس وفاعليته.

كما تهدف هذه الدراسة لإيجاد المؤشرات الفاعلة للكشف عن أحادية البعد بصورة واضحة وبالتالي خدمة صانعي الاختبارات وتسهيل مهمتهم في استخدام المؤشر الأكثر فاعلية في الكشف عن مدى تحقق افتراض أحادية البعد ووضع الاختبارات المناسبة للأهداف المناسبة. علاوة على ذلك تقدم هذه الدراسة نموذجا جديدا باستخدام النظرية البنائية والتحليل العاملي التوكيدي للكشف عن أحادية البعد باسلوب جديد، وهذا يعكس الأهمية النظرية للدراسة.

أما الأهمية العملية لهذه الدراسة فتنبع في إمكانية إيجاد مؤشرات فاعلة تكشف عن افتراض أحادية البعد بحيث تمكن معدي الاختبارات من بناء مقاييس تتمتع ببعد واحد قادرة على قياس القدرة الواحدة. وهذا الأمر يخدم المعلمين كثيرا في قطاع وزارة التربية والتعليم وخاصة في اختبارات القدرة المتنوعة والاختبارات التشخيصية لأغراض الكفاية المختلفة، ولإعداد بنوك اختبارات تتمتع بخصائص سيكومترية جيدة من شأنها استخدام صور متعددة للمقياس الواحد.

التعريفات الاصطلاحية والإجرائية:

فاعلية المؤشرات: مدى مطابقة المؤشرات المستخدمة في الكشف عن افتراض أحادية البعد ضمن حدود القطع المتفق عليها في الأدب النظري، ويتم مقارنة المؤشرات معا تبعا للاختبار الإحصائي مربع كاي لحسن المطابقة عند حدود القطع في حساب الباقي المعياري (أقل من -2 وأكبر من 2).

افتراض أحادية البعد: وجود قدرة واحدة فقط تفسر أداء الفرد في اختبار ما وتتميز فقرات الاختبار طبقاً لذلك بأنها لا تختلف فيما بينها إلا من حيث مستوى صعوبتها، بمعنى انه إذا كانت التوزيعات المشروطة لدرجات الاختبار عند مستوى قدرة معين لمجتمعات فرعية عدة متطابقة فان الاختبار يكون أحادي البعد. أما إذا كان هناك اختلاف ملحوظ بين هذه التوزيعات فإنه يوجد قدرة أخرى تؤثر في الأداء.

التحليل العاملي التوكيدي: تقنية إحصائية تتيح الفرصة لتحديد واختبار صحة نماذج معينة للقياس، وتتمثل الإجراءات المتبعة فيها في تحديد النموذج المفترض والذي يتكون من المتغيرات الكامنة، وفي ضوء افتراض التطابق بين مصفوفة التغاير (Covariance Matrix) للمتغيرات الداخلة في التحليل والمصفوفة المفترضة من قبل النموذج، ينتج العديد من المؤشرات على جودة هذه المطابقة والتي يتم من خلالها قبول النموذج المفترض للبيانات أو رفضه في ضوئها، وتعرف بمؤشرات جودة المطابقة.

افتراضات الدراسة:

1. ثبات المتغير الأساسي وهو افتراض أحادية البعد، بمعنى إن افتراض أو درجة أحادية البعد لا تختلف باختلاف طول الاختبار، ولا تختلف باختلاف شكل توزيع القدرة، ولا تختلف باختلاف النموذج المستخدم.
2. تكرار عملية التحليل الإحصائي لاستجابات المفحوصين على الفقرات لخمس مرات متتالية (5 replication) لكافة متغيرات الدراسة؛ من حيث طول الاختبار، وشكل توزيع القدرة، والنموذج المستخدم. لضمان ثبات القيم (Stability Replication)، وللحصول على مؤشرات تتمتع بالثبات ليتم اعتمادها بالشكل النهائي.

محددات الدراسة:

يمكن بيان محددات الدراسة بما يلي:

1. اقتصرت الدراسة على البيانات المولدة نظرا لحاجة الدراسة لأعداد كبيرة من المفحوصين، ولعدم إمكانية توفرها تم الاعتماد عليها في إجراء التحليلات الإحصائية المختلفة.
2. اقتصرت الدراسة على 1000 مفحوص لجميع مستويات الدراسة، على اعتبار أن عدد المفحوصين متغير مضبوط لجميع متغيرات الدراسة.
3. اقتصرت الدراسة على برمجية AMOS 21 في التحليلات الإحصائية الخاصة.
4. اقتصرت الدراسة الحالية على البيانات المولدة من النماذج اللوجستية الأحادية البعد.

الفصل الثاني

الإطار النظري والدراسات السابقة

الإطار النظري للدراسة:

تشير أهمية الدراسة إلى أن فاعلية المؤشرات المتبعة في الكشف عن افتراض أحادية البعد تتطلب المعرفة بالأساس النظري الذي ترتكز عليه، حيث يشير كل من هامبلتون وروفينيل (Hambleton & Rovinelli, 1986) إلى ضرورة تعريف مصطلح "أحادية البعد " (Unidimensionality) وأساليب قياسه للتحقق من توفره أو غيابه في مجموعة من فقرات الاختبار. وعلى ذلك؛ يثار سؤالان:

الأول: ما المقصود بمصطلح أحادية البعد؟

الثاني: ما الأسلوب المناسب للتحقق من هذا الافتراض لكي يكون متسقا مع التعريف؟

وقد بين هامبلتون وروفينيل (Hambleton & Rovinelli, 1986) أن مصطلح أحادية البعد هو مفهوم مجرد ليس له تعريف إجرائي محدد، إلا أنه يقصد به – بصفة عامة – أن هناك سمة واحدة فقط في الاختبار يمكن أن تفسر أداء الفرد. وتفترض بعض نماذج السمات الكامنة كما يشير علام (2005) إلى وجود قدرة واحدة فقط يمكن أن تفسر أداء الفرد في اختبار ما وتتميز فقرات الاختبار طبقاً لهذه النماذج بأنها لا تختلف فيما بينها إلا من حيث مستوى صعوبتها، بمعنى أن تدرج صعوبة فقرات المقياس السلوكي بحيث تعرف فيما بينها متغيرا واحدا.

ويذكر أن افتراض أحادية البعد يصعب تحقيقه بشكل تام، فهناك العديد من العوامل التي تؤثر في استجابة الأفراد كالدافعية والقلق والناحية المعرفية وغيرها، ويتطلب تحقيق هذا

الافتراض وجود عامل مسيطر واحد على الاختبار وهــو القــدرة المقاســة (Anastasi & Urbina ,1997).

وقبل الحديث عن طرق الكشف عن أحادية البعد، لا بد من الإشــارة إلــى بعــض المفاهيم والمصطلحات المتعلقة بأحادية البعد، والتطرق إلى الآثار الناجمــة عــن انتهــاك افتراض أحادية البعد، والتعرف إلى الأسس النظرية التي تربطها بنظرية الاستجابة للفقــرة (Item Response Theory – IRT) الواجب توافرها للتحقق من ذلك.

نظرية استجابة الفقرة (Item Response Theory)

تعرف نظرية الاستجابة للفقرة في القياس باسم السمات الكامنة لاهتمامها بــالربط بين استجابة الفرد لفقرة اختبار ذات خصائص معينة وقدرته (الشافعي، 2008)، كما أنهــا تركز على تحديد مواقع الأفراد على المقاييس النفسية والتربوية، وانبثقت عن هذه النظرية مجموعة من النماذج؛ تهدف جميعها لتحديد العلاقة بين أداء الفرد في الاختبــارات، وبــين السمات أو القدرات، وتوصف من خلال الدالة الرياضية لكل نموذج من هذه النماذج لتقدير موقع الفرد في البناء أو السمة الكامنة، ومن ثم تفــسير أدائــه أو التنبــؤ بأدائــه اللاحــق (Gustafsson, 1980).

وتعد نماذج استجابة الفقرة دوال رياضية احتمالية تصف منحنيات الفقــرات فــي الاختبار. بافتراض توافر مجموعة كبيرة من الفقرات التي تقيس السمة نفسها، بحيث يكون تقدير القدرة مستقلاً عن عينة الفقرات المستخدمة في القياس، وكذلك بافتراض وجود مجتمع كبير من الأفراد، كما يمكن تقدير الخصائص السيكومترية للفقرات (الــصعوبة والتميــز) مستقلاً عن عينة الأفراد التي استخدمت في تقدير هذه الخصائص (علام ،2005).

خصائص نماذج السمات الكامنة:

يشير هامبلتون (Hambleton, & Swaminathan, 1985) إلى أن نظريـــة الـــسمات الكامنة ترتكز على مسلمتين أساسيتين هما:

- إن أداء الأفراد في الاختبار يمكن التنبؤ به (يفسر) من خلال مجموعـــة عوامـــل تسمى سمات او قدرات.
- ان العلاقة بين أداء الأفراد على الفقرة ومجموعة السمات المفترض إنها تؤثر في الأداء على الفقرة نفسها يمكن ان توصف بواسطة الدالة التزايدية التـــي تـــسمى دالـــة خصائص الفقرة (Item Characteristic Function) وهذه الدالة تحدد الأفراد الذين لهم درجات مرتفعة في القدرات او السمات والذين يملكون احتمـــالات متوقعـــة مرتفعـــة للإجابة الصحيحة على الفقرة، من الأفراد الذين لهم درجات منخفضة في القـــدرات أو السمات.

وتستند هذه النماذج على عدد من الافتراضات التي يجب أن تتوافر فـــي البيانـــات المستمدة من الاختبارات، مستفيدة من ثلاث معالم هي تمييز الفقرة وصعوبة الفقرة والخـــط التقاربي الأسفل (Lower- Asymptote) أو التخمين وهو منـــاظر لاحتماليـــة الاســـتجابة الصحيحة التي قد تحدث بالصدفة.

وتقيس بعض الاختبارات قدرة واحدة، وأخرى قدرات متعددة، ففي بعـــض نمـــاذج نظرية استجابة الفقرة Item Response Theory يفترض وجود قدرة واحدة تفسر أداء الأفراد على الاختيار، وهذا ما يسمى بنماذج أحاديـــة البعـــد Unidimensional Models، وبعـــض النماذج تفترض وجود أكثر من قدرة تفسر أداء الفرد على الاختبار تدعى بنماذج متعـــددة الأبعاد Multidimensional Models. ومن طرق تصنيف النماذج أحادية البعد هو تصنيفها تبعاً لاستجابات الأفراد، ومستويات الاستجابة الشائعة هي: ثنائية التدريج (Dichotomous)

ومتعـددة التـدريج (Polytomous) واسـتجابات متـصلة (Continuous)، (Reckase & Mckinley, 1991; Reckase, 1979).

وقد أشارت كثير من الأدبيـات (Ackreman, 1994; Embretson & Reise, 2000; Knol & Berger, 1991) إلى أن النماذج متعددة الأبعاد يمكن التمييز بينها كونهـا نمـاذج استكشافية تعويـضية جزئيـاً (Exploratory Models partially Compensatory Models) ونماذج توكيدية غيـر تعويـضية جزئيـا Confirmatory Models Non Compensatory Models، وقد بين اكرمان (Ackerman, 1994) أنه في النماذج التعويضية يستطيع الأفـراد الضعاف في قدرة معينة أو أكثر تعويض ضعفهم جزئياً، بعبارة أخرى تسمح هذه النمـاذج بالتفاعل بين قدرات متعددة الأبعاد بحيث النقص في إحداها يعوض فـي قـدرات أخـرى مرتفعة، ومن الأمثلة على النماذج التعويضية النموذج اللوجستي ثنائي المعلمـة للفقـرات الثنائية الذي طوره ريكس (Reckase, 1985).

أما في حالة النماذج غير التعويضية ففيها لا تعوض قدرة الفرد بالاستجابة عن سمة مختلفة، حيث أن الفرد يعتمد في نماذج الاستجابة للفقرة متعددة الأبعاد على أكثر من قـدرة للإجابة عن فقرات الاختبار، وفي بعض الأحيان تتطلب فقرات الاختبار مركبا لعـدد مـن المهارات أو القدرات لكي يجيب إجابة صحيحة، أي أن نقص إحدى القـدرات لا يعـوض تعويضا كاملا بالقدرات الأخرى المرتفعة، وتعد النماذج متعـددة الأبعـاد غيـر شـائعة الاستخدام وذلك لان متطلبات تطبيقها كثيرة لتقدير معالم الفقـرة والأفـراد بـشكل ملائـم (Embretson & Reise, 2000).

وتعددت نماذج استجابة الفقرة الأحادية والمتعددة التدريج Polytomous Item Response Models and Unidimesionality، ومن أشهرها (Embretson & Reise, 2000):

- نموذج الاستجابة المتدرجة: (Graded Response Model)(GRM): قدم هذا النموذج سميجيما (Samejima)، ويعتبر هذا النموذج تعميما للنموذج ثنائي المعلم، ويستخدم فقرات ليس بالضرورة أن تكون متساوية في عدد فئات الاستجابة عليها، حيث أن عدم تحقق هذا الشرط لا ينشا عنه أي تعقيدات في تقدير معالم الفقرة أو تفسيرها.

- نموذج الاستجابة المتدرجة المعدل: (Modified Graded Response Model): طور هذا النموذج موراكي (Muraky) كتعديل لنموذج الاستجابة المتدرجة، بحيث يسهل استخدامه في تحليل الاستجابات على فقرات مقياس التقدير (Rating Scale) كما في فقرات الاتجاهات حيث يكون لجميع الفقرات العدد نفسه من فئات الاستجابة، لذا يعد هذا النموذج المعدل لنموذج الاستجابة المتدرجة، حالة خاصة منه.

- نموذج التقدير الجزئي: ((Partial Credit Model (PCM)): يعد نموذج التقدير الجزئي توسيعا لنموذج راش المتعلق بالفقرات ثنائية التدريج، حيث تم تطويره ليصبح بالإمكان تقدير معالم الفقرات التي يمكن أن تأخذ في عملية التصحيح أكثر من قيمتين، وبشكل يعكس القدرة التي تقيسها الفقرة.

- نموذج التقدير الجزئي المعمم: (Generalized Partial Credit Model): طور موراكي (Muraki) نموذج التقدير الجزئي المعمم الذي يسمح باختلاف معلم الميل للفقرات ضمن المقياس الواحد وأطلق عليه نموذج التقدير الجزئي المعمم.

كما تعددت نماذج السمات الكامنة الأحادية البعد الثنائية التدريج وتباينت من حيث أشكال المنحنيات المميزة تبعاً لاختلاف عدد (بارامترات) أو معالم الفقرات وهي (Hulin, Drasgow, & Parsons, 1983) :

- النموذج أحادي المعلم (نموذج راش) Rasch Model: ويسمى هذا النمــوذج بنمــوذج الترجيح أللوغاريتمي أحادي البارامتر IPL وذلك لأنه يستخدم الدالة الآســية فــي التنبــؤ بالاحتمالات، وتشمل على بارا متر (معلم) واحد فقط، وهو معلم صعوبة الفقرات لتمثيــل الفروق الفردية بين الفقرات.

و يعتبر من أبسط النماذج وأكثرها شهرة، والدالة الرياضية التي تعبر عنه تربط بين احتمال إجابة الفرد إجابة صحيحة عن الفقرة ومعلمة صعوبة الفقرة، ومنحنيات خــصائص الفقرة لهذا النموذج تختلف فقط في موقعها على متصل السمة .كما يفترض أن جميع الفقرات متماثلة في قدرتها التمييزية وأن معلمة التخمين لها تساوي صفراً.

- النموذج ثنائي المعلم (نموذج بيرنبوم) Birnbaum Model: وهذا النموذج يسمح للفقرات بأن تختلف في كل من معلم الصعوبة ومعلم التمييز. ولم يأخذ بالاعتبــار معلــم التخمــين كعامل مؤثر على أداء الفرد على الاختبار، حيث يفترض أن الطلبة الضعاف ذوي القــدرة المنخفضة لم يمكنهم التخمين من الإجابة عن الفقرات الصعبة إجابة صحيحة، لذلك يفترض هذا النموذج أن معلم التخمين يساوي صفرا، وأن منحنيات خصائص الفقرة تختلف في ميلها كما تختلف في موقعها على متصل، لذلك تضمنت الصيغة الرياضية لهذا النمــوذج معلــم تمييز الفقرة و معلم الصعوبة فقط.

- النموذج ثلاثي المعلم (نموذج لورد) Lord Model: يسمى هذا النموذج اللوغارتمي ثلاثي المعلمة (3PL)Three – Parameter Logistic Model) إذ أضاف معلما ثالثا فــي احتمــال توصل الأفراد للإجابة الصحيحة وهو معلم التخمين، ومنحنيات خــصائص الفقــرة لهــذه النماذج تختلف في ميلها (معلم التمييز)، وموقعها على متصل السمة (معلم الصعوبة)، وخط التقارب السفلي لمنحنى خصائص الفقرة (معلم التخمين)، وبذلك يعتبر أكثر النماذج عمومية

لأنه يسمح باختلاف معالم الفقرة الثلاثة، وبهذا يجب ان يؤخذ ذلك بعين الاعتبار عند مطابقة البيانات المستمدة من الاختبار لهذا النموذج.

وقد عبر عنها بيكر (Baker, 2001) من خلال المعادلة الآتية:

$$P_i(\theta) = c_i + (1-c_i)\left[\frac{1}{1+e^{-1.7a(\theta-bi)}}\right]$$

حيث أن:

θ: قدرة الفرد.

$P_i(\theta)$: احتمال أن الفرد الذي يمتلك القدرة θ أجاب على الفقرة i إجابة صحيحة.

b_i : معلمة صعوبة الفقرة.

a_i : ميل دالة الفقرة عند النقطة b ، وتشير إلى معلمة تمييز الفقرة .

C_i : خط التقارب السفلي لدالة الفقرة، وهي تشير إلى معلمة التخمين.

افتراضات نظرية السمات الكامنة الأحادية البعد Assumptions of Latent Trait Theory:

ترتكز نماذج السمات الكامنة إلى مجموعة من الافتراضات يجب توافرها في البيانات المستمدة من الاختبار، ومجموعة الافتراضات هي: أحادية البعد حيث يقيس الاختبار سمة واحدة فقط، والاستقلال الموضعي وهو استقلال أداء الفرد على فقرة الاختبار عن أدائه على فقرة أخرى من الاختبار نفسه، منحنى خصائص الفقرة ويمثل مخطط بياني لاحتمالية الإجابة الصحيحة على الفقرة، كما تفترض معظم النماذج المستعملة أن عامل السرعة لا يؤدي دورا في الإجابة عن الفقرة وأن إخفاق الأفراد في الإجابة عن فقرات الاختبار يرجع إلى انخفاض قدراتهم وليس لتأثير عامل السرعة في إجاباتهم، واللاتباين ويعني أن معالم الفقرة لا تعتمد على التوزيع الإحصائي للسمة أو الخاصية المراد قياسها، وأن المعالم التي تصف أداء الأفراد

لا تعتمد على فقرات الاختبار، وفيما يأتي توضـــيحا لهــذه الافتراضـــات: Anastasi & .Crocker & algina, 1986; Hambelton & Jones, 1993; Lord, 1980;) Urbina, 1997

افتراض أحادية البعد Unidimesionality

إن افتراض أحادية البعد شرط أساسي لبعض نماذج استجابة الفقرة. وهذا الشرط مطلوب في اختبارات القدرة وبعض الاختبارات التحصيلية والعديد من الاختبارات النفسية والتربوية، وهي تفترض وجود قدرة واحدة فقط يمكن أن تفسر أداء الفرد في اختبار ما وتتميز فقرات الاختبار طبقاً لهذه النماذج بأنها لا تختلف فيما بينها إلا من حيث مستوى صعوبتها، بمعنى أن تدرج صعوبة فقرات المقياس السلوكي بحيث تعرف فيما بينها متغيراً واحداً. أما النماذج التي تفترض وجود أكثر من قدرة واحدة تكمن وراء أداء الفرد على الاختبار فتسمى نماذج متعددة الأبعاد، ويمكن أن يكون الاختبار أحادي البعد في مجتمع ومتعدد الأبعاد في مجتمع آخر. فإذا كان اختبار معين متحيزا ثقافيا بشكل كبير فان هذا الاختبار يمكن أن يكون أحادي البعد لكل المجتمعات التي لها نفس الخلفية الثقافية ويتحيز لها الاختبار، ولكن عندما يطبق الاختبار على مجتمعات بثقافات مختلفة يمكن أن يمتلك الاختبار أكثر من بعد لتفسير الأداء على هذا الاختبار.

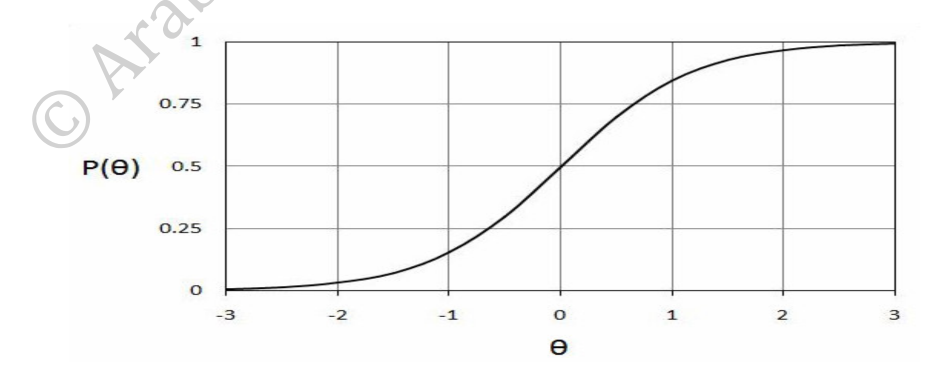
افتراض الاستقلال الموضعي (Local Independence)

أن هذا الافتراض يعني الاستقلال الإحصائي بين الإجابات لمختلف فقرات الاختبار لشخص معين، وكذلك الاستقلال الإحصائي بـــين الإجابـــات لمختلـــف الأشـــخاص أي أن الاستجابات السابقة لشخص معين على فقرات الاختبار ليس لها تأثير في الاستجابات على بقية فقرات الاختبار نفسه، وأن استجابات شخص معين في اختبار ما لا تؤثر في استجابات أي شخص أخر في ذلك الاختبار، وقد ذكر لورد (Lord, 1980) أن الاستقلال المركزي ينتج

تلقائياً من أحادية البعد، أي بتحقيق فرض الاستقلال المركزي إذا كانت مجموعــة فقــرات الاختبار تقيس قدرة أو سمة واحدة وإذا تبين أن الاستجابات على المقياس أحادية البعد فانه يمكن افتراض أن فرض الاستقلال المركزي قد تحقق، وعندما يكون احتمال الحصول على أي نمط معين لدرجات فرد ما مساوياً حاصل ضرب احتمالات حدوث هذا النمط لكل فقرة من فقرات الاختبار التي أجاب عنها.

افتراض منحنى خصائص الفقرة (Item Characteristic Curve)

منحنى خصائص الفقرة هو تمثيل بياني لاحتمالية الإجابة الصحيحة على الفقــرة، وهي دالة متزايدة بحيث يزداد احتمال إجابة الفرد عن الفقرة إجابة صحيحة بزيادة قدرتــه التي تعبر عن مقدار السمة لديه. ويمثل المحور الأفقي (قدرة الفرد) والمحــور الرأســي (احتمال الإجابة الصحيحة للفرد على الفقرة). ويتم رسم المنحنى بواسطة دوال رياضــية من البيانات الأولية المستخدمة في منحني انحدار الفقرة على الاختبار، وتكون هذه الــدوال مختلفة باختلاف نماذج السمات الكامنة اعتمادا على مجموعة متنوعة مــن الافتراضــات. ويوضح الشكل 1 هذا المنحنى.



الشكل 1:رسم توضيحي لمنحنى خصائص الفقرة

حيث أن: θ قدرة الفرد، $P(\theta)$:احتمال الإجابة الصحيحة للفرد على الفقرة.

فبعض هذه النماذج يستخدم دوال المنحنى التراكمي ألاعتدالي، والــبعض الأخــر يستخدم الدوال اللوغارتمية التراكمية، والتي تفيد في بعض الخصائص الرياضية للعلاقــات اللوغارتمية. وبشكل عام فإن شكل المنحنى لخصائص الفقرة يأخــذ شــكل الحــرف (S- Shaped) ويتم وصف منحنى خصائص الفقرة بعدد من المعالم يختلف عــددها بــاختلاف نموذج الاستجابة للفقرة، ومن الشائع أن يكون عددها واحدا أو اثنين أو ثلاثة. ويعبر التغير في ميل المنحنى عن تمييز الفقرة، كما يعبر عن تغير موقعها على متصل الــسمة بمعلــم الصعوبة، في حين أن خط التقارب السفلي لمنحنى خصائص الفقــرة يــشير إلــى معلــم التخمين. وإن الفقرة ذات مستوى صعوبة مناسب تمثله نقطة على ميزان القــدرة، حيــث يكون احتمال الإجابة الصحيحة (0.5) (بشرط عدم تأثر الإجابة بالتخمين العــشوائي)، وهي نقطة انقلاب المنحنى (Inflection Point). وإذا كانت الفقرة أكثر صعوبة فإن المنحنى تحدث له إزاحة إلى اليمين دون أي تغيير في شكله، أما إذا كانت الفقرة أقل صعوبة، فــإن الإزاحة تكون إلى اليسار.

افتراض عامل السرعة في الإجابة Speedness

إن هذا الافتراض نادراً ما يشار إليه لكونه متضمناً في افتراض أحادية البعد بمعنى انه عندما يؤثر عامل السرعة في الأداء على الاختبار تكون هنالك ســمتان علــى الأقــل تؤثران في الأداء، هما سرعة الأداء والسمة التي يقيسها الاختبار. وبهذا لا يتوافر افتراض أحادية البعد في الاختبار، ويمكن تقدير مدى تأثير عامل السرعة في الإجابة عــن طريــق معرفة عدد المفحوصين الذين لم يكملوا إجابة جميع فقرات الاختبار، الذي طبق عليهم ومن ثم فإن تحديد الوقت المناسب للإجابة هو الذي يبعد المفحوصين عن اللجوء إلى التخمين.

اللاتباين لمعالم الفقرة Invariance Of Item Parameter

هو الوصف التقني لاستقلالية القياس التي تعبر عن شروط الموضوعية والدقة في القياس النفسي ومن المفاهيم المهمة فيها، وهذا يعني أن معالم الفقرة لا تختلف عندما تحسب من مجتمعات مختلفة، أي أنها توفر مميزات القياس ذي الفئات المتساوية، أي أنه يمكن اختبار المجموعات والأفراد بفقرات مختلفة تناسب مستويات قدراتهم وتكون قابلة للمقارنة المباشرة. فدرجة الفرد لا تعتمد فقط على عدد الفقرات التي أجاب عليها بصورة صحيحة وإنما على مستويات الصعوبة المحسوبة مسبقاً لتلك الفقرات أيضا.

ويفسر اللاتباين هذا بأن منحنى خصائص الفقرة يمثل دوال انحدار، إذ تعد القدرة متغيراً مستقلاً والاستجابة على الفقرة متغيراً تابعاً. وتتميز دوال الانحدار بأنها لا تتغير بتغير التوزيع للمتغير المتنبئ، لذا فأن احتمالية الاستجابة الصحيحة عند مستوى قدرة معين يعتمد فقط على القدرة لا على عدد الأفراد. حيث إن خطوط الانحدار لا تتباين فإن الخط ألتقاربي الأدنى ونقطة الانعطاف والميل عند نقطة ما (قدرة أو سمة) تبقى نفسها بغض النظر عن توزيع القدرة في المجموعة. لذا فإن معالم الفقرة لا تتباين بل تبقى نفسها بغض النظر عن اختلاف مجموعة المفحوصين.

ولهذا فإن لكل مستوى قدرة توجد احتمالية واحدة للإجابة الصحيحة تتمثل بمنحنى خصائص الفقرة، وهذه الاحتمالية مستقلة عن توزيع قدرات الأفراد في مجتمع المفحوصين، أي أنها لا تعتمد ولا ينبغي أن تعتمد على عدد المفحوصين في اي مجموعة عند مستويات القدرة المختلفة لذلك فإن المنحنى المميز للفقرة يطبق بشكل متساوي على مستويات القدرة جميعها.

وبسبب خاصية اللاتباين للمعالم هذه، يمكن حساب معالم الفقرات من خلال بيانات مجموعة واحدة من المفحوصين، ومن ثم تطبيقها وبثقة على مجموعات أخرى بما فيها

المجتمع الكلي. وهذه الميزة تجعل لنماذج السمات الكامنة تطبيقات واسعة في مجالات مثل الاختبارات التكيفية المفصلة وبنوك الأسئلة، واكتشاف تحيز الفقرات، ومجالات أخرى عديدة.

القياس المتحرر من أثر صعوبة الفقرات ومن أثر قدرات الأفراد /Item – Free Person Person –free Measurment

لتحقيق الموضوعية في النماذج المختلفة لا بد من تحقق شرط تحرر القياس من أثر صعوبة الفقرات، وتحرر القياس من مجموعة الفقرات المستخدمة،وهذا يعني ثبات تقدير كل من قدرة الفرد وصعوبة الفقرة واستقرارها بالرغم من اختلاف مجموعة الفقرات المستخدمة في القياس، طالما أنها عينة ملائمة، وطالما أن هذه المجموعات المختلفة من الفقرات تقع على ميزان تدريج واحد، أي أنها تعرف متغيرا واحدا. كما يتم بناء تقدير معلمات فقرات الاختبار المتحرر من تأثير العينة، بمعنى آخر استقلال تدريج الفقرات عن عينة الأفراد (Hambelton and Swaminathan, 1985).

دالة معلومات الفقرة والاختبار Item (tests) Information Function

وهي دالة تمثل العلاقة ما بين متغيرين هما قدرة الفرد والمعلومات المقدمة من خلال هذه الفقرات، وتعبر هذه الدالة عن كمية المعلومات التي تقدمها الفقرة عند مستوى القدرة التي تقيسها، وذلك بتحديد أقصى ارتفاع لمنحنى دالة معلومات الفقرة عند مستوى معين للقدرة. وإن كمية المعلومات المستمدة من الفقرة تتغير بحسب مستوى القدرة، فكلما كانت الفقرة واقعة على نقطة أعلى على منحنى خصائص الفقرة كانت المعلومات أكثر.

وتستخدم دالة المعلومات للتأكد من دقة تقدير معالم الفقرة ومعلم القدرة، وتأتي أهمية دالة معلومات الفقرة من كون مساهمة كل فقرة باقتران معلومات الاختبار يتحدد بشكل مستقل عن باقي فقرات الاختبار. وبشكل عام فان الفقرات ذات التمييز العالي تساهم بقوة اكبر في تأكيد

دقة القياس من تلك ذات التمييز المتدني، حيث تعطي الفقرة أفضل مساهمة لها في تأكيد دقة القياس حول قيمة صعوبتها (b) على متصل القدرة، أي أن دالة المعلومات تعتمد على ميل اقتران الاستجابة للفقرة والتباين عند كل مستوى للقدرة (θ). فكلما زاد الميل فإن التباين يقل وتزداد المعلومات التي توفرها الفقرة. فإذا ما كان المنحنى مزاحا باتجاه اليمين فهذا يعني أن الفقرة صعبة، وإذا كان ارتفاعه كبيرا فهذا يعني أن تمييز الفقرة عال والعكس صحيح. وتختلف معادلة منحنى معلومات الفقرة حسب النموذج اللوجستي كما يأتي (Baker, 2001) :

- النموذج اللوجستي أحادي المعلم وتعطى معادلة منحنى معلومات الفقرة بالعلاقة التالية:

$$I_i(\theta) = P_i(\theta)\, Q_i(\theta)$$

حيث أن:

$I_i(\theta)$: دالة المعلومات للفقرة (i).

$P_i(\theta)$: احتمالية الاستجابة الصحيحة على الفقرة (i).

$Q_i(\theta)$: احتمالية الاستجابة الخاطئة على الفقرة (i).

- النموذج اللوجستي ثنائي المعلم فان معادلة منحنى معلومات الفقرة تعطى بالعلاقة التالية :

$$I_i(\theta) = a_i^2 P_i(\theta)\, Q_i(\theta)$$

حيث أن:

a_i^2 : مربع معلم التمييز للفقرة (i) .

– النموذج اللوجستي ثلاثي المعلم فان معادلة منحنى معلومات الفقرة تعطى بالعلاقة التالية

$$I_i(\theta) = a^2\left[\frac{Q_i(\theta)}{P_i(\theta)}\right]\left[\frac{P_i(\theta)-c^2}{(1-c^2)}\right]$$

حيث أن: c_i^2: مربع معلم التخمين للفقرة (i).

ويرى هامبلتون وسوامنثيان (Hambelton and Swaminathan, 1985) انه كلما زاد عدد الفقرات تزداد قيمة معلومات الاختبار، بالإضافة إلى انــه كلمــا زاد الميــل تــزداد المعلومات التي توفرها الفقرة وهذا يعني أن الفقرات التي تتضمن معالم تمييز كبيرة تقــدم معلومات اكبر عن قدرة الأفراد وبالتالي الحصول على دقة اكبر، وعليه يمكن انتقاء فقرات اختبار اعتمادا على كمية المعلومات التي تسهم بها الفقرات في كميــة المعلومــات الكليــة للاختبار، و يمكن التعبير عن دالة معلومات الاختبار بالصيغة الآتية:

$$I(\theta) = \sum I_i(\theta)$$

حيث أن:

$I_i(\theta)$: اقتران معلومات الفقرة (i)، $I(\theta)$: اقتران معلومات الاختبار.

ولأن الاختبار مكون من مجموعة من الفقرات؛ فإن المعلومات التي يقدمها الاختبار عند أي مستوى قدرة هي مجموع مقدار معلومات الفقرات عند ذلك المــستوى . لــذا فــان المستوى العام لاقتران معلومات الاختبار سوف يكون أكبر من المعلومات لفقرة مفرد ة.

ويشير كروكر و ألجينا (Crocker & Algina, 1986) أن اقتران المعلومات له فوائد متعددة مثل مقارنة اختبارات مختلفة تقيس السمة نفسها، أو مقارنــة طــريقتين مختلفتــين لتصحيح نفس الاختبار، بالإضافة إلى مقارنة استراتيجيات مختلقة في تــشكيل اختبــارات مختلفة مثل المقارنة بين الاختبار التقليدي والاختبار التكيفي وذلك من خلال كمية المعلومات التي يعطيها كل اختبار منها.

كما يشير هامبلتون وسوامينثان (Hambleton & Swaminathan, 1985) إلى أن كمية المعلومات التي تعطيها مجموعة من الفقرات عند مستوى قدرة معين تتناسب عكسياً مع الخطأ المعياري في تقدير القدرة، فتزود مطور الاختبار بمستوى الخطأ المعياري للاختبار عند كل مستوى قدرة. ولذلك فان الاختبار الذي تم تحليل فقراته من الممكن حساب دالة معلوماته من خلال جمع دوال معلومات الفقرات المكونة له. فإذا كانت كمية المعلومات عالية فان الخطأ في التقدير يكون منخفضا، بمعنى أن العلاقة عكسية بينهما:

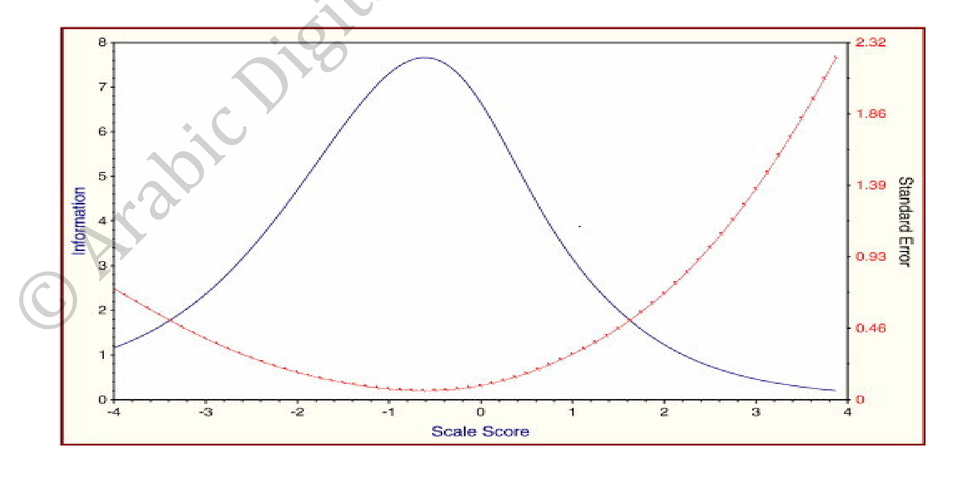
$$SEE(\theta)=1/[I(\theta)]^{1/2}$$

حيث أن:

SEE (θ): الخطأ المعياري في التقدير .

I(θ): اقتران المعلومات للاختبار

ويمكن التعبير عنها من خلال الشكل 2:



الشكل 2 : رسم توضيحي لمنحنى اقتران معلومات الاختبار

الآثار الناجمة عن انتهاك افتراض أحادية البعد:

لقد شدد هـامبلتون وسـواميناثان (Hambleton & Swaminathan, 1985) علـى وجوب تحقق افتراض أحادية البعد عند تطبيق نماذج نظرية استجابة الفقرة أحادية البعد، إذ أن انتهاك هذا الافتراض يؤثر بصورة واضحة عند تطبيـق المعادلـة الخاصـة بحـساب ارجحية نمط استجابة المفحوصين.

وأكد هيولن ودراسغو ووبارسـونز(Hulin, Drasgow, & Parsons, 1983) علـى ضرورة التحقق من افتراض أحادية البعد، وعند التحقق من ذلك يمكن تطبيـق المعـادلات الخاصة لتقدير معالم الفقرات والمفحوصين. ولكن من الناحية العملية قد لا يتحقق افتراض أحادية البعد في بعض المواقف الاختبارية، فقد يحتوي الاختبار على مجموعة من الفقرات تقيس سمة معينة ويحتوي أيضا على مجموعة فقرات تقيس أخرى، وهكذا. وبالتالي يتكون الاختبار من مجموعة من الاختبارات الفعلية يتحقق في كل منها فرض أحادية البعد، وعندئذ يتم إيجاد معالم الفقرات ومعالم قدرات المفحوصين في كل اختبار على انفراد. لذلك لا بـد من التحقق من هذا الافتراض قبل استخدام النماذج أحادية البعد.

ويوجد العديد من الاختبارات التي تقيس أكثر من سمة، وبـذلك ينتهـك افتـراض أحادية البعد، إذ يختلف المفحوصون في مستوى قدراتهم في إحدى السمات. لذا يجب ان لا تفترض أحادية البعد على البيانات، بل يجب أن يجري التحقـق منهـا مـسبقا باسـتخدام المؤشرات المختلفة. وقد أظهرت بعض الدراسات أن استخدام نماذج نظرية استجابة الفقرة أحادية البعد لبيانات اختبارات متعددة الأبعاد، يعمل على التقليـل مـن مطابقـة البيانـات للنموذج، مما يؤدي إلى نتائج غير صحيحة عن طبيعة البيانات (Ackerman, 1994).

وبين ووكر وبرتراس (Walker & Beretras, 2000) أن تطبيق بيانـــات متعـــددة الأبعاد في نموذج يفترض أحادية البعد يؤدي إلى زيادة في خطأ القياس، وبالتالي يقود لنتائج غير صحيحة حول معالم الفقرة، ومعلم القدرة للأفراد، ويعطي تقديرا مرتفعا للثبات ودالـــة معلومات الاختبار. كما بين أن استخدام نموذج استجابة الفقرة أحادي البعد علـــى بيانـــات متعددة الأبعاد يعد استنتاجا غير صحيح لتوزيع قدرات المفحوصين على الأبعاد الثانويـــة. وأوضحت دراسة وانج ووانج (Wang & Wang, 2003) أن عدم تحقق افتـــراض أحاديـــة البعد يؤثر بصورة واضحة في تقدير المعالم المختلفة، وبالتالي على دقة هذه التقديرات.

ونظراً لأهمية الموضوع فقد طرح هاتي (Hattie, 1985) خمس مؤشرات تـــستخدم للكشف عن مدى توفر شرط أحادية البعد هي: نمط الاستجابة Answer Patterns ومعـــاني الثبات Reliability، والمكونات الرئيسة Principal Components، والتحليل العاملي Factor Analysis، ومطابقة البيانات لأحد نماذج السمات الكامنة Latent Trait Models. وأظهـــر من خلالها أن العديد من المؤشرات ذات مسوّغ ضعيف، وبعـــضها الآخـــر عـــدل نتيجـــة الانتقادات الموجهة إليه.

وفيما يلي عرض مختصر لهذه المؤشرات:

نمط الاستجابة Answer Patterns

ويتحقق ذلك بإتباع نموذج (جوتمان) A Guttman Response Pattern في الاستجابة على الفقرات من خلال ترتيب الفقرات بحسب صعوبتها. فالفرد الذي يخفق في الاســـتجابة على الفقرة بمستوى صعوبة معين يخفق في الاستجابة على فقرة بمـــستوى صـــعوبة أعلى.

الاتساق الداخلي أو طريقة الثبات Reliability Methods

ويقاس من خلال ارتباط الفقرة بالدرجة الكلية الذي يعني أن الفقرة صادقة في قياس ما يجب قياسه، فمعامل الارتباط هنا يعبر عن صدق الفقرة، وبذلك يكون الاختبار صادقاً في قياس ما وضع لأجله. فالفقرات ذات معامل الارتباط الأعلى (r = 0.3)، تستبقى في الاختبار لان مثل هذه الفقرات عند تجميعها تؤلف تدرجاً بأعلى اتساق داخلي، وهذا يعطي مؤشراً على أن مجموعة الفقرات تقيس سمة أو قدرة واحدة. أما مؤشر الثبات المعتمد فهو مؤشر كرونباخ الفا، وهذا المؤشر دليل على الاتساق الداخلي للفقرات وهذا يشير إلى أحادية البعد (Nunnally , 1979) .

تحليل المكونات الرئيسة (PCA) Principal Component Analysis

وتقوم فكرة تحليل المكونات الأساسية PCA على إنشاء مجموعات خطية (Linear Combinations) يطلق عليها المكونات الرئيسية يكون الهدف منها شرح التباين المشترك بين هذه المكونات باستخدام عدد من الفقرات للحد من البيانات أو تخفيضها. والغرض من ذلك هو تقييم الأبعاد وتحسين تفسيرها، ومنها:

1- نسبة التباين المفسر للعامل الأول: أشار ريكاس (Reckase, 1979) إلى أنه إذا وصلت نسبة التباين المفسر للعامل الأول إلى (% 20) من التباين الكلي دل ذلك على أحادية البعد، وبذلك يشرح المكون الرئيس الأول أعلى تباين. وهذا التباين يعبر عنه عادة كنسبة من التباين الكلي، وتحذف الفقرات التي لا تنسجم مع العامل العام Common Factor. وبتكرار العملية يمكن التوصل إلى الفقرات التي تقيس العامل العام فقط. في حين يرى كارمينز وزلير (Carmines & Zeller, 1979) أن النسبة يجب أن تصل إلى (% 40) كحد أدنى للإشارة إلى أحادية البعد.

2- نسبة الجذر الكامن الأول إلى الجذر الكامن الثـــاني (The Ratio of the Eigen values Associated with the first Principal Component ($\lambda1$) and the second Principal Component ($\lambda2$)): وذلك بحساب الجذور الكامنة لتحقيق أحادية البعـد لمجموعــة فقرات الاختبار بأن تكون قيمة الجذر الأول كبيرة مقارنة بالجذر الثاني. فقد أشـــار جلورفيلد (Glorfeld, 1995) إلى أنه يجب أن تكون قيمة الجذر الكامن الأول أكبـــر من واحد($\lambda1 >1$)، والثاني قريباً في قيمته من الجذور الأخرى وأقل من واحد ليـــدل على أن الفقرات أحادية البعد. إن نسبة الجذر الكامن الأول إلى الجذر الكامن الثـــاني يعطي مؤشرا معقولا على افتراض أحادية البعد، حيث تتحدد النسبة بالمعادلة الآتية:

$$\text{Ratio} = \lambda1 / \lambda2$$

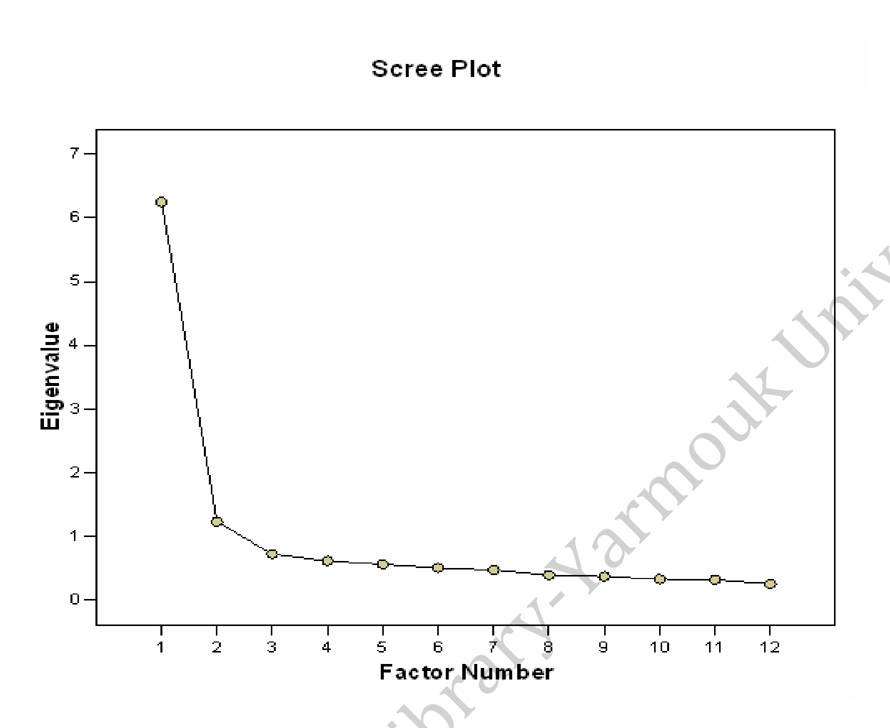
حيث أن:

(λ_1): الجذر الكامن للمكون الأول، (λ_2): الجذر الكامن للمكون الثاني.

واقترح هاتي (Hattie, 1985) أن النسبة المرتفعة للجذر الكامن الأول إلى الجذر الكامن الثاني تشير إلى أحادية البعد والقيم الدنيا منها تشير إلى تعدد الأبعـــاد. وحـــددت تلـــك النسبة بحيث تكون أكبر أو تساوي 2 تبعا لبرنامج Winsteps manual، للدلالـــة علـــى أحادية البعد (Linacre, 2008).

3- التمثيل البياني لقيم الجذور الكامنة (مخطط سكري) Eigenvalues Scree Plot :

يعتمد على رسم خاص للجذور الكامنة يوضح من خلاله المكونات الداخلة في التحليـــل ومنه يمكن الحكم على الجذور الأكثر تفسيرا للتباين المشترك. فعندما يصبح الخط شبه أفقي يتم التوقف عنده، والشكل الآتي مثال توضيحي فقط، حيث يحتفظ فقـــط بالعامـــل الأول والعامل الثاني .



الشكل 3: تمثيل بياني يوضح قيم الجذور الكامنة

مطابقة البيانات لأحد نماذج السمات الكامنة:

يوجد مؤشرات عديدة تعتمد على نماذج السمة الكامنة أحادية المعلم وثنائية المعلم، ويمكن اعتماد افتراض الاستقلال الموضعي لكل زوج من الفقرات الاختبارية ولأي مجتمع متجانس من المفحوصين كمؤشر لأحادية البعد إذ يترتب على الاستقلال المركزي وجـود سمة واحدة يقيسها الاختبار. وتقاس جودة الفقرة بمدى ملائمة درجتها للنمـوذج، وذلـك باستخدام إحصاءات الملاءمة، وهي من نوع إحصاءات ملائمة الفقرة الكلية Total Fit Item Statistics، ويتخصص بتحديد مدى ملاءمة الفقرة بوجـه عـام لافتراضـات النمـوذج، وإحصاءات الملاءمة للفقرة بين المجموعات Between Fit Item Statistics بتحديـد مـدى استقرار المستوى النسبي لصعوبة الفقرات عبر مستويات القدرة حتى يتم الوصـول إلـى جودة الاختبار، وأيضاً (إحصاء ملاءمة الفـرد Person-FitStatistics) الـذي يتخـصص

بتحديد مدى ملاءمة كل فرد من الأفراد للنموذج. و يوجد نوعان آخران مـن إحـصاءات الملاءمة هما إحصاء الملاءمة المتقاربة In Fit وإحـصاء الملاءمـة المتباعـدة Out Fit (Greene, 1993).

إن المشكلة الرئيسة في تقييم افتراض أحادية البعد، تكمن في اختلاط هذا المصطلح بشكل متبادل مع المصطلحات الأخرى كالثبات والاتساق الداخلي والتجانس. فهنالك فـرق واضح بين هذه المصطلحات، وخلاصة القول لا يوجد مؤشر ثابت للكشف عـن مؤشـر أحادية البعد يزود بقرار واضح، لذا جاءت هذه الدراسة لتكشف عن افتراض أحادية البعـد بمؤشرات مختلفة (Crocker & Algina, 1986; Hattie, 1985).

التحليل العاملي Factor Analysis :

عند الحديث عن أسلوب التحليل العاملي Factor Analysis لا بد من وجود فرضية رئيسية وهي أن العوامل الرياضية تمثل متغيرات كامنة Latent Variables تدعى بالأبعاد السيكولوجية التي يمكن تحديد طبيعتها فقط عن طريق فحص طبيعة المتغيرات التي لهـا إحداثيات مرتفعة المحاور المتعامدة. والهدف من إجراء التحليل العـاملي تمييـز الأبعـاد المحدودة التي يفترض أن تشكل العديد المهام بطرق كمية. وتعتبر العوامل Factors التـي ينتجها أسلوب التحليل العاملي متغيرات رياضية ينظر إليها على أنها محاور تصنيفية يمكن من خلالها أن تتجمع المعلومات المكتسبة من الاختبارات المختلفة. وكلما كانت قيمة المحور للمتغير التي يشار إليها بتشبع المتغير Loading على العامل كبيرة كلمـا ازدادت أهميـة العامل في التأثير على العلاقات بين ذلك المتغير ومجموعة المتغيرات الأخـرى (Hattie, 1985).

والمفهوم الإحصائي للتحليل العاملي ويعني العملية التي يتم فيها تلخيص مجموعـة كبيرة نسبيا من المتغيرات المترابطة بأقل عدد ممكن من المتغيرات غير المترابطة. أما المفهوم

النفسي للتحليل العاملي فهو البحث عن العوامل التي تؤثر في الظواهر المعقدة التي تـرتبط بعضها بعض. ويمكن عن طريق هذا المنهج تحديد العوامل المسؤولة عن السلوك؛ فبدلاً من تعددها وتشعبها تصبح محدودة نسبياً لان التحليل العاملي يهدف إلى التبسيط العلمي والـى إيجاد مجموعة من القدرات أو السمات الأقل في العدد والاشمل في طبيعتها من الاختبارات بغية الوصول إلى اختبارات نقية. ولكن هذه العوامل المستخلصة بإحدى طرائـق التحليـل العاملي يصعب تفسيرها نفسياً إلا بعد تدويرها وتبسيطها قدر الإمكان (Glorfeld, 1995).

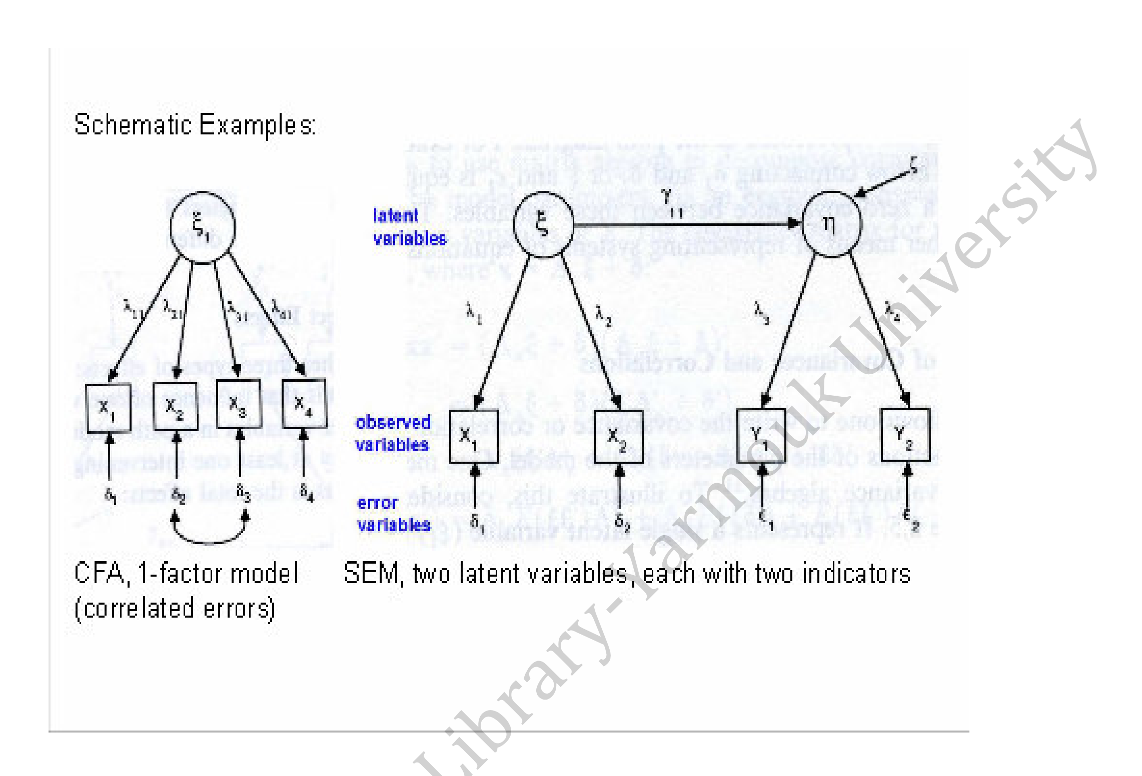
وينظر إلى العوامل على أنها محاور يتم تدويرها بهـدف جعـل العلاقـات بـين المتغيرات وبعض هذه العوامل أقوى ما يمكن. ويوجد عدة طرق لتدوير المحـاور، مـن أكثرها شيوعًا هي طريقة تعظيم التباين المعروفة باسم Varimax وهي طريقة تدوير تتميز بأنها تحافظ على خاصية الاستقلال بين العوامل. وهذا يعني هندسيًا بقاء المحـاور أثنـاء عملية التدوير متعامدة. وهناك طريقة أخرى تسمى Quartimax تسمح بوجود عوامل مائلة مترابطة. ويكون الغرض من أي عملية تدوير هو الحصول على صورة لقيم التشبع على العوامل لإنتاج مصفوفة العوامل لإجراء التحليلات الإحصائية الخاصـة وتحديـد الأبعـاد وتصنيف البيانات بشكل مناسب. ويطلق على هذا الأسلوب: أسـلوب بالتحليـل العـاملي الاستكشافي Exploratory Factor Analysis (Stevens, 2002).

وهناك أسلوب آخر يطلق عليه التحليل العاملي التوكيـدي Confirmatory Factor Analysis الذي يمكن الباحث من تحديد عدد العوامل مسبقًا ووضع فرضيات تتعلق بنمط قيم التشبع للمتغيرات. وعادة يستخدم للحد من خطأ القياس من خلال عـدد مـن المؤشـرات (مؤشرات جودة المطابقة) لكل متغير، والقدرة على اختبار النماذج لإختيار النموذج الأنسب من بين عدد من النماذج البديلة لمطابقة البيانات للنماذج المقترحة (Bollen, 1990).

ويعتبر التحليل العاملي التوكيدي أحد تطبيقات نموذج المعادلة البنائية (Structural Equation Modeling (SEM))، الذي يتيح الفرصة لتحديد واختبار صحة وصدق نماذج معينة للقياس يتم بناؤها في ضوء أسس نظرية سابقة متفق عليها. وتتمثل الإجراءات المتبعة في أسلوب التحليل العاملي التوكيدي في تحديد النموذج المقترح مسبقاً في ضوء أطر نظرية، وهو النموذج البنائي المفترض والذي يتكون من (Timothy, 2006):

1. المتغيرات الكامنة (Latent Variable)، المنوي دراستها في النموذج.
2. المتغيرات الخارجية (Exogenous Variable) أو المتغيرات المستقلة: وتمثل الأبعاد المفترضة للمقياس وهي تعامل بشروط ثابتة، وتعتبر مستقلة عن أخطاء القياس، ومنها تخرج أسهماً متجهة إلى النوع الثاني من المتغيرات.
3. المتغيرات التابعة أو المتغيرات الداخلية (Endogenous Variable): والتي تمثل العبارات الخاصة بكل بعد أو الأبعاد الخاصة بكل عامل عام ويتم تحديد قيمها في النموذج، وهنا يفترض أن العبارات مؤشرات للمتغيرات الكامنة.

وعادة يتم رسم مخطط هيكلي مقترح يربط بين المتغيرات السابقة موضحا العلاقات فيما بينها كما هو الحال في الشكل 4:



الشكل 4:رسم توضيحي لمخطط هيكلي مقترح في التحليل العاملي التوكيدي

يوضح الشكل 4 مخططين لنموذجين، النموذج الأول في اليسار يعتبر مـن أبـسط النماذج الهيكلية فهو يربط متغيراً كامناً واحداً بأربعة متغيرات مستقلة أو خارجية، ويـربط النموذج الهيكلي الثاني بين متغيرين كامنين بمتغيرين مستقلين لكل واحد منهما، في حين أن الكثير من الدراسات والأبحاث تتعدد بها المتغيرات كما تتعدد بها العلاقات السببية وبالتالي يصبح المخطط أكثر تعقيدا.

ولا تختلف الإجراءات المتبعة في تحليل الانحدار أو تحليل المسار باستخدام نموذج المعادلة البنائية والتي قد يطلق عليها نماذج المعادلة الآنية كثيراً عما هو متبع فـي حالـة التحليل العاملي التوكيدي الذي من شأنه توضيح العلاقات السببية بين المتغيرات في النموذج.

ففي تحليل المسار يتم تحديد قيمة معاملات المسارات والذي يدل على الأثر المباشر لعلاقة السبب بالنتيجة، وذلك من خلال البناء التخطيطي لأشكال النماذج السببية المختلفة. وتستخدم الكلمات " السبب "و " النتيجة: لتناظر المتغيرات الخارجية والمتغيـرات الداخليـة علـى الترتيب، وهذا يعني أنه أسلوب إحصائي ينظم الارتباطات الداخلية للنموذج ليوضح الأثـر والتأثر بين تلك المتغيرات عن طريق فحص معاملات المسار بينها لتحديد فيما إذا كانـت هناك علاقة سببية (Bollen, 1990).

أما في التحليل العاملي التوكيدي فيتم حساب كافة الارتباطـات بـين المتغيـرات الداخلية والخارجية ويتم تشكيل مصفوفة معاملات الارتباط التي من خلالها يـتم حـساب مصفوفة التغاير. وفي ضوء افتراض التطابق بين مصفوفة التغاير للمتغيرات الداخلة فـي التحليل والمصفوفة المفترضة(المقترحة) من قبل النموذج، ينتج العديد من المؤشرات على جودة هذه المطابقة؛ وعلى أساسها يتم قبول النموذج المفترض للبيانات أو رفضه. وتعرف تلك المؤشرات بمؤشرات جودة المطابقة Indicator Fit Index (Greene, 1993).

نموذج المعادلة البنائية:

ويعد موضوع نمذجة المعادلة البنائية من المواضيع الواسعة التي تظهر في معظـم المجالات النفسية والاقتصادية، وقد كتب عنها العديد من العلماء فـي أبحـاثهم المختلفـة والمشار إليهم في بولن (Bollen, 1990) أمثال (Duncan, 1975; Greene, 1993; Fox, 1984; Bollen, 1989)، وهي عائلة من التقنيات الإحصائية تضم وتشمل تحليل المسار والتحليـل العاملي التوكيدي وتتميز باحتوائها على عدد كبير من المؤشرات تكشف عـن المتغيـرات الكامنة، والأسهم التي تربط المتغيرات معا هي نوع من التحليل العاملي التوكيـدي، تقـوم أساسا على التباينات المشتركة فيما بينها والذي هو محور نموذج المعادلة البنائية.

وينهج نموذج المعادلة البنائية مناحي عدة في الإجراءات التحليلية التوكيدية وهي (Bentler, 1987):

1. المنحى التوكيدي الصارم (Strictly Confirmatory Approach): وهو نهج فيه نوع من التشدد والصرامة يستخدم جودة المطابقة Goodness-of-Fit Tests باستخدام إجراءات المعادلة البنائية لتحديد ان كان نمط الفروق بين التباينات والتغايرات للبيانات ثابت ويتفق مع نموذج المسار الذي تم اختياره من قبل الباحث.
2. منحى النماذج البديلة (Alternative models approach): وتستخدم لفحص أثنين أو أكثر من النماذج السببية (Causal Model)، وهنا يتم اختبار النموذج الأكثر تطابقا مع البيانات، بحيث يتم بناء أكثر من نموذج سببي لمعرفة مدى مطابقتها للبيانات وهو من الطرق المفضلة، ولكن المشكلة التي يقع بها الباحث أنه لا يوجد في الأدب النظري أكثر من نموذجين متطورين بديلين للمفاضلة بينهما.
3. منحى النماذج المتطورة (Model development approach): وعادة تجمع النماذج البنائية المتطورة بين الأغراض الاستكشافية و التوكيدية Confirmatory and Exploratory Purposes، وهنا يتم بناء نموذج باستخدام إجراءات المعادلة البنائية لاختباره وفقا لطرق جودة المطابقة، واختيار نموذجا آخر بديل يعتمد على بعض الافتراضات للاعتماد عليه في بعض الحالات، وهو النهج الأكثر شيوعا في الأدب النظري، وتخضع هذه النماذج لطرق جودة المطابقة للبيانات، ويحق للباحث في هذا المنحى استخدام بيانات معايرة Calibration Data للتحقق من الطرق التوكيدية.

أهمية المعادلة البنائية:

يلجأ إليها العديد من الباحثين في بحوثهم وذلك (Anglim, 2007):

1. لتقدير العلاقات بين المتغيرات الكامنة (Latent Variables) .
2. تسمح في اختبار وفحص النماذج التنافسية (Competing Models) .
3. في الاستكشاف المباشر وغير المباشر لمجموعة المتغيرات الداخلة للنموذج.
4. استكشاف العلاقات المتعددة بين المتغيرات بطرق تكاملية (An Integrated Manner).
5. تستخدم في تطوير المقاييس المختلفة (Scale Development)، ويعتبر التحليل العاملي التوكيدي أحد تطبيقاتها، وفحص النماذج المعقدة والنمـــاذج البـــسيطة (Simple and Complex Meditational Models).

يؤكد بلنتر (Bentler, 1987) أن النماذج البنائية تتطلب عينات كبيرة للتأكد من جودة مطابقتها لعينة البيانات المستخدمة، إذ أن العينات الصغيرة لا تقدم دلالات إحصائية واضحة مما يوقع الباحث في مشاكل أثناء فحص جودة المطابقة.

ويلعب التحليل العاملي التوكيدي دوراً هاماً وبارزاً في نموذج المعادلة البنائية؛ من خلال تقييم دور خطأ القياس، وصحة بناء النموذج Validating the Measurement Model. فقد بين كلاين (Kline, 1998) أنه لا بد من توليد وبناء النموذج بطرق رياضية صـــحيحة تعتمد على أسس نظرية صحيحة، ومن ثم اختبار هذا النموذج وفقا للتحليلات الإحـــصائية المتبعة لنموذج المعادلة البنائية، باستخراج عدد من المؤشرات لجميع المتغيرات، وأخيـــراً اعتماد النموذج الأنسب للبيانات.

وتخضع النماذج البنائية لتحليلات إحصائية معقدة، تشمل علاقات إحصائية معقـــدة بين المتغيرات الملاحظة والمتغيرات غير الملاحظة الكامنة. كما هو الحال في تحليل متعدد

الانحدار وتحليل المسار، ويضطر الباحث فيها لبناء نماذج مختلفة لإجراء عمليات جـودة المطابقة، ولا يستطيع الباحث القيام بتلك العمليات الحسابية المعقدة بنفسه ممـا يـضطره اللجوء لبرمجيات خاصة متطورة، للحصول على مؤشرات جودة المطابقة مثـل LISREL (Linear Structural Relations), AMOS (Analysis of Moment Structures), and EQS (Environmental Quality Standard)، وجميعها تؤدي الهدف نفسه مـع بعـض التحفظـات للشركات المنتجة لها، حيث أن برمجية LISREL, AMOS من إصدارات أل SPSS ويعتمد عليها معظم الباحثين في مجالات العلوم المختلفة. إلا أن برمجيـة (AMOS (Analysis of Moment Structures) تقدم عدد كبير من المؤشرات قـد تـصل إلـى 25 مؤشـر، فـي حين(LISREL (Linear Structural Relations يقدم 15 مؤشرا، وعادة تقدم إصدارات SPSS ثلاثة نتائج لنماذج مختلفة يمكن إجمالها فيما يلي (Kline , 1998):

1. النموذج المشبع (The Saturated Model): وهذا النوع استكـشافي بحـت Fully Explanatory، تأخذ فيه غالبية قيم مؤشرات جودة المطابقة الواحد الصحيح فأكثر، وعلى هذا الأساس فإن بعـض المؤشـرات مثـل RMR(Root Mean Square) Residual، RMSEA (Root Mean Square Error of Approximation) لا يمكـن حسابها، وبالتالي تكون هذه النماذج ضعيفة في تقديرها وحكمها على صحة بنـاء النموذج.
2. النموذج المستقل The Independence Model: أو ما يسمى بالنموذج الصفري The Null Model والذي يتم بناؤه بشكل افتراضي من قبل البرمجية الإحصائية ذاتهـا، وفيه تكون المتغيرات غير مترابطة وهذا يعني أن معاملات الارتباط جميعها قـيم صفرية، أي بمعنى أن جميع قيم المؤشـرات سـتكون غيـر دالـة، وأن بعـض المؤشرات (GFI) تعتمد على البيانات وبالتالي لا يمكن حسابها.

3. النموذج الافتراضي The Default Model: أو نموذج الباحث المقترح وهو أفضل النماذج مقارنة بالنماذج الاستقلالية والنماذج المشبعة، يتم بناؤه من قبل الباحث ويدخل فيه العلاقات بين المتغيرات جميعها ويتم حساب كافة الارتباطات بينها، وتكون فيه مؤشرات جودة المطابقة أكثر قوة وتفسيرا لأنه يعتمد على البيانات بشكل مباشر، لذا يعتمد عادة على نتائج النماذج الافتراضية في تفسير نتائج مؤشرات جودة المطابقة للحكم على صحة النموذج.

افتراضات نموذج المعادلة البنائية:

يشير بولن (Bollen ,1989) إلى ضرورة توافر عدد من الافتراضات الواجبة في البيانات والمتغيرات لإجراء نموذج المعادلة البنائية وهي:

1. التوزيع الطبيعي المتعدد للمؤشرات Multivariate Normal Distribution of the Indicators: ينبغي على كل قيمة من قيم مؤشرات جودة المطابقة أن تتمتع بتوزيع طبيعي، وان الخروج عن هذا الافتراض قد يتسبب باختلافات كبيرة بقيم المؤشرات، وبشكل عام يعتبر انتهاك هذا الافتراض مشكلة كبيرة، فقد يتسبب بتضخيم قيم مؤشر مربع كاي the Chi-Square Test، فمن المعلوم أن التوزيع الطبيعي مطلب أساسي في أسلوب الأرجحية العظمى للتقدير، والتي هي محور نموذج المعادلة البنائية في تقدير المعاملات، وبالتالي انتهاك هذا الافتراض سينعكس على تقديرات الأرجحية العظمى، كما أن أخطاء القياس تزداد بانتهاك افتراض التوزيع الطبيعي، وعلى الرغم من أهمية هذا الافتراض إلا أن الدراسات والأدب النظري (Reise & Waller, 2003; Hambelton, Swaminathan, & Rogers ,1991) لا يبدي أي اهتمام بهذا الافتراض.

2. التوزيع الطبيعي المتعدد للمتغيرات التابعة الكامنة Multivariate Normal Distribution of the Latent Dependent Variables: يجب أن تتمتع جميع قيم المتغير التابع في النموذج بتوزيع طبيعي، وأن المتغيرات الثنائية التدريج الكامنة تتسبب بانتهاك هذا الافتراض، ولمثل هذه الحالات يوصي العديد من الباحثين باستخدام التحليل الطبقي للتخلص من مشكلة المتغيرات الثنائية التي لا تتناسب مع نموذج المعادلة البنائية.

3. الخطية Linearity: يفترض نموذج المعادلة البنائية وجود علاقة خطية بين المتغيرات الكامنة والمؤشرات وبين المتغيرات أنفسها، وان انتهاك هذا الافتراض سوف يتسبب بتحيز في تقديرات المطابقة وأخطاء القياس لصالح النموذج، ويرى كلاين (Kline, 1998) أن النماذج غير الخطية تساهم في انتهاك افتراض التوزيع الطبيعي، كما أنه يجب أن لا يتم إغفال مصفوفة التغاير والتي تفترض الخطية في حساباتها.

4. القياس غير المباشر Indirect Measurement: عادة جميع المتغيرات في نموذج المعادلة البنائية هي متغيرات كامنة، لا يتم حسابها بصورة مباشرة.

5. المؤشرات المتعددة Multiple Indicators: يتوافر في نموذج المعادلة البنائية عدد كبير من المؤشرات لقياس المتغيرات الكامنة، كما يسمح بإجراء مطابقة بين النماذج المختلفة، على عكس تحليل الانحدار الذي يعتبر حالة خاصة من نموذج المعادلة البنائية والذي يستخدم مؤشر واحد للمتغيرات الكامنة.

وبناءً على ما سبق فقد ركز هذا البحث على أسلوب التحليل العاملي التوكيدي كأحد تطبيقات نموذج المعادلة البنائية، حيث يتم من خلاله الحصول على عدد من المؤشرات

ويطلق عليها بمؤشرات جودة المطابقة (لمطابقة البيانات مع النموذج المتوقع)، وسيتم عرضها بنوع من التفصيل وبيان أهمية كل مؤشر وكيفية قراءته.

مؤشرات جودة المطابقة:

وتأتي هذه المؤشرات على هذا النحو (Timothy, 2006):

- النسبة بين χ^2 ودرجات الحرية المطلقة df.
- مؤشرات المطابقة المطلقة: Absolute Fit Indexes
- مؤشرات المطابقة المتزايدة: Incremental Fit Indexes

ويمكن تفصيلها كما يلي:

النسبة بين قيمة χ^2 ودرجات الحرية df :

يؤكد هوبر وكوهلان ومولن (Hooper, Coughlan & Mullen, 2008) انه إذا كانت النسبة تتراوح بين 5 و 2 دل على قبول النموذج بشكل ضعيف، ولكن إذا كانت أقل من 2 دل على أن النموذج مطابقا بصورة كبيرة للبيانات، وهناك العديد من الدراسات والبحوث (Bentler & Chou ,1987) التي تستخدم دلالة χ^2 كمؤشر لجودة المطابقة، وهذا مقبول في حالة العينات كبيرة الحجم (التي تزيد عن 400) أو عندما لا يكون رغبة في مقارنة نماذج بنائية مختلفة لنفس البيانات، حيث تتأثر هذه القيمة بحجم العينة ولذلك لابد وأن تؤخذ في الاعتبار بعض المؤشرات الأخرى لجودة المطابقة بجانب النسبة بين قيمة χ^2 ودرجات الحرية df، في حين أن بايرن (Byrne, 1998) يعتبر النسبة ($ratio > 2.00$) مؤشر على عدم مطابقة النموذج للبيانات.

ولغرض إجراء اختبار جودة التطابق والملاءمة للنماذج المقترحة لبيانات الدراسة تستخدم الصيغة التالية (Bentler & Bonett, 1980):

$$(\chi^2 / df) = \sum\left(\frac{(Oi - Ei)^2}{Ei}\right) / df$$

حيث:

Oi : القيم الملاحظة Observed Values

Ei : القيم المتوقعة Expected Values

df : درجات الحرية degree of freedom

مؤشرات المطابقة المطلقة (Absolute Fit Indexes) :

ويكون الاهتمام هنا منصباً على مقارنة مصفوفة التغاير للعينة (Covariance Matrix) وهي المصفوفة الأساسية بالمصفوفة المحللة، وتشمل المؤشرات الآتية: (Bentler & Bonett, 1980).

1 – مؤشر حسن المطابقة (GFI) Goodness of Fit Index:

ويقيس هذا المؤشر مقدار التباين في المصفوفة المحللة (البيانات الفعلية) عن طريق النموذج المقترح موضوع الدراسة، وهو بذلك يناظر مربع معامل الارتباط المتعدد في تحليل الانحدار المتعدد وتتراوح قيمته بين (0 ، 1) وتشير القيمة المرتفعة بين هذا المدى إلى تطابق أفضل للنموذج مع بيانات العينة وقد طور هذا المؤشر ليتحرر من تعقيد النموذج ويعرف المؤشر الجديد بمؤشر حسن المطابقة المعدل لدرجات الحرية Adjusted Goodness of Fit Index (AGFI).

ويعطى بالمعادلات الآتية:

$$GFI = 1 - (V_{residual} / V_{total})$$

حيث أن:

$V_{residual}$: تباين البواقي في مصفوفة التغاير.

V_{total}: التباين الكلي في مصفوفة التغاير.

$$AGFI = 1 - (1 - GFI)\frac{d_b}{d},$$

حيث d درجات الحرية للنموذج، d_b درجات الحرية للنمــوذج الــصفري (Null Model) والنموذج الصفري هو النموذج الذي تم بناءه تبعا للفرضية الصفرية، وفيه تكــون المتغيرات الكامنة غير مترابطة، ويطلق عليه أيــضا النمــوذج المــستقل Independence Model وهو متواجد بصورة افتراضية by default داخل البرمجيات والحزم الإحــصائية المستخدمة في التحليل.

2- مؤشر حسن المطابقة المتشدد (PGFI : parsimony goodness of fit index):

وهو مؤشر معدل من GFI بدرجات الحرية وهو يتشدد كثيرا في حساباته لأنه يشير إلى عدد المعلمات المقدرة اللازمة لتحقيق مستوى معين لمطابقة النموذج المقترح، لذا تكون قيمته المحسوبة أقل من قيم المؤشر GFI، إلا أنها تبقى محصورة ضمن المدى (0 ، 1) وتشير القيمة المرتفعة بين هذا المدى إلى تطابق أفضل للنموذج مع بيانات العينة ويعطى بالمعادلة الآتية (Mulaik, et al. 1989):

$$PGFI = GFI\frac{d}{d_b}$$

3- مؤشر جذر مربعات الأخطاء Root Mean Square Residual (RMR): يعد من مؤشرات جودة المطابقة، وهو الجذر التربيعي لمتوسط مربع البواقي لمصفوفة التباين المشترك للنموذج الملاحظ وللنموذج الضمني (Observed Covariance and Model-Implied Covariance)، وكلما كانت قيمته قريبة من الصفر دل على أفضل تطابق، والقيم المقبولة تتراوح بين (0.05-0.08)، وإن كانت قيمته أقل من 0.05 يعطي تطابقا أكثر للبيانات، ويعطى بالمعادلة الآتية:

$$RMR = \sqrt{\sum_{g=1}^{G}\left\{\sum_{i=1}^{p_g}\sum_{j=1}^{j\leq i}\left(\hat{s}_{ij}^{(g)} - \sigma_{ij}^{(g)}\right)\right\} \Big/ \sum_{g=1}^{G} p^{*(g)}}$$

حيث أن:

σ_{ij}: التباين المشترك لمصفوفة النموذج المتوقع Implied Covariance .

$\check{s}_{ij}$: التباين المشترك لمصفوفة النموذج الملاحظ Observed Covariance.

i, j: تشير إلى عدد البيانات الداخلة في التحليل في كلا المصفوفتين .

$d_b = \sum_{g=1}^{G} p^{*(g)}$: درجات الحرية للنموذج الصفري. ويشير الرمز g إلى عدد المجموعات الداخلة في التحليل (g = 1, 2, …, G).

4- مؤشر جذر متوسط مربع الخطأ التقريبي Root Mean square Error of Approximation (RMSEA):وهو من أهم مؤشرات جودة المطابقة لأنه يقيس التباين لكل درجات الحرية، ويعتبر أكثر ملاءمة للعينات الكبيرة التي تزيد عن 200 مفحوص، وكلما صغرت قيمة هذا المؤشر دل على التطابق الأكبر، فإن كانت قيمته 0.05 فأقل دل ذلك

على أن النموذج يطابق البيانات وإذا كانت القيمة محصورة بين(0.05 ، 0.08) دل ذلــك على أن النموذج مطابق بدرجة معينة لبيانات العينة أما إذا زادت قيمته عن 0.08 فيــتم رفض النموذج. وكلما كانت اقرب الى الصفر دل على التطابق التام، وتعطــى المعادلــة على النحو الآتي (Bentler & Bonett, 1980):

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(F_i - O_i)^2}$$

حيث أن:

Fi: القيم المتوقعة Expected Value

Oi: القيم الملاحظة او المشاهدة Observation Value

N: عدد القيم المتوقعة والملاحظة.

مؤشرات المطابقة المتزايدة Incremental Fit Indexes:

وهي تعتمد في تقديرها على مقارنة النموذج المقرح مع النموذج الصفري والــذي يُفترض فيه وجود عامل عام واحد تتشبع عليه كــل المتغيــرات المقاســة ويــشمل هــذه المؤشرات (Bentler & Bonett, 1980):

1- مؤشر المطابقة المعياري (NFI) Normed Fit Index: يتم من خلاله مقارنة النموذج المقترح بالنموذج الصفري، وتتراوح قيمته بين (0، 1) وتشير القيمة المرتفعة (أكثر من 0.95) إلى تطابق أفضل للنموذج مع البيانات. ويعطى بالعلاقة الآتية:

$$NFI = \frac{\chi^2(\text{Null Model}) - \chi^2(\text{Proposed Model})}{\chi^2(\text{Null Model})}$$

Null Model: النموذج الصفري الذي تم بناءه وفقا للفرضية الصفرية by default.

Proposed Model: النموذج المقترح من قبل الباحث.

وتعتبر القيم المقبولة بين 0.90 و 0.95، في حين ان كانت أكبر من 0.95 هي مطابقة بصورة كبيرة للنموذج مع البيانات، أما أقل من 0.90 تعتبر اقل مطابقة. ومما يجدر الإشارة إليه أنه يتأثر بحجم الاختبار، فكلما زاد طول الاختبار تقل قيمة هذا المؤشر، واشتق منه مؤشر جديد Parsimony Adjustment to the NFI: PNFI وهو معدل لمؤشر المطابقة المعياري بدرجات الحرية للنموذج الصفري والنموذج المتوقع وهو أكثر تشددا لدرجات الحرية ويعطى بالمعادلة (James, Mulaik & Brett, 1982; Mulaik, et al., 1989):

$$PNFI = NFI\ (d / d_b)$$

حيث أن d درجات الحرية للنموذج المتوقع، d_b درجات الحرية للنموذج الصفري.

2- مؤشر المطابقة النسبي (Relative fit index (RFI: وهو من المؤشرات المعدلة لدرجات الحرية للمؤشر NFI. وتشير القيم المرتفعة (أكبر من 0.90) إلى تطابق البيانات مع النموذج أما ان كانت أكبر من 0.95 تشير الى أفضل تطابق. وتشير القيمة واحد صحيح إلى التطابق التام.

3- مؤشر المطابقة المقارن (Comparative Fit Index (CFI: وتتراوح قيمة هذا المؤشر بين (0، 1) وتشير القيمة المرتفعة بين هذا المدى إلى تطابق أفضل للنموذج مع بيانات العينة. وتعطى بالمعادلة الآتية:

$$CFI = \frac{d(\text{Null Model}) - d(\text{Proposed Model})}{d(\text{Null Model})}$$

حيث أن:

$$d = \chi^2 - df$$

df: degrees of freedom درجات الحرية للنموذج.

إذا كانت قيمته اكبر من واحد فإنها توضع واحد صحيح وان قلت عن الصفر أيضا توضع صفر، ودائما ما تكون قيمة CFI أكبر من قيمة TLI، لارتباط هذين المؤشرين معا بصورة كبيرة. ويوجد منه مؤشر جديد معدل لمؤشر المطابقة المقارن Parsimony Adjustment to the CFI: وهو معدل بدرجات حرية وأكثر تشددا ويعطى بالمعادلة الآتية (James, Mulaik & Brett, 1982; Mulaik, et al., 1989):

$$PCFI = CFI\ (d / d_b)$$

حيث أن d درجات الحرية للنموذج المتوقع، d_b درجات الحرية للنموذج الصفري.

4- نسبة برسموني The Parsimony Ratio: وهي نسبة درجات الحرية للنموذج المتوقع والنموذج الصفري، وتدخل قيمته في حساب المؤشرات المتشددة Parsimony Index وبالتالي يؤثر على قيم المؤشرات الأخرى ولا يستخدم كمؤشر منفرد للحكم على جودة النموذج، وهو يتأثر كثيرا بدرجات الحرية فكلما زادت درجات الحرية أعطى قيما أعلى وبالتالي تتأثر قيمة المؤشر المحسوب منه، ويعطى بالمعادلة الآتية (James, Mulaik & Brett, 1982; Mulaik, et al., 1989):

$$\mathbf{PRATIO} = \frac{d}{d_i}$$

حيث أن d درجات الحرية للنموذج المتوقع، d_I درجات الحرية للنموذج الصفري.

5- مؤشر توكر لويس Tucker-Lewis Index (TLI) or Non-Normed Fit Index NNFI :

ويطلق عليه أيضا مؤشر المطابقة غير المعياري NNFI ، وتتراوح قيمة هذا المؤشر بين (0 ، 1) وتشير القيم المرتفعة من هذا المدى إلى تطابق أفضل للنموذج مع بيانات العينة. ويعطى بالمعادلة الآتية (Kenny, Kaniskan, and McCoach , 2011):

$$TLI = \frac{\chi^2/df(\text{Null Model}) - \chi^2/df(\text{Proposed Model})}{\chi^2/df(\text{Null Model}) - 1}$$

إذا كانت قيمته أكبر من واحد صحيح توضع واحد صحيح، يلاحظ ان هذا المؤشر يعتمد بشكل أساسي على مؤشر χ^2/df فإن لم تتغير هذه النسبة فإن مؤشر (TLI) لن يتغير، ومن الجدير بالذكر أن قيم كل من TLI و CFI تتأثر نسبياً بحجم العينة، إذ أنها تعطي قيما أقل بارتفاع حجم العينة (Bentler & Bonett, 1980).

6- مؤشر المطابقة المتزايد Incremental Fit Index (IFI): وتتراوح قيمة هذا المؤشـــر بين (0، 1) وعادة تشير القيمة المرتفعة منه (أكبر من 0.9) إلى تطابق أفضل للنمـــوذج مع بيانات العينة. وتعطى بالمعادلة الآتية:

$$IFI = \frac{\chi^2(\text{Null Model}) - \chi^2(\text{Proposed Model})}{\chi^2(\text{Null Model}) - df(\text{Proposed Model})}$$

مؤشر هولتر لجودة المطابقة Hoelter's (1983)`critical N' for a significance level of .05: يقارن هولتر (Hoelter's ,1983) قيمة المؤشر في التحليل بالحجم المسموح به وهــو يناظر حجم العينة في التحليل الإحصائي المستخدم، فإن كانت قيمة مؤشر هولتر أكبر مــن الحجم المسموح به فهذا يعني تطابق البيانات للنموذج، وان كان أقل يتم رفــض النمــوذج وتكون المقارنة عادة على مستوى الدلالة ($\alpha = 0.05, 0.01$). وهو يرتبط بــصورة كبيــرة بمؤشر χ^2 ويعطى بالمعادلة الآتية (Kenny, Kaniskan, and McCoach, 2011)

$$\text{Hoelter Index} = \frac{[1.645 + \sqrt{(2df - 1)}]^2 + 1}{2\chi^2/(N - 1) + 1}$$

حيث أن df درجات الحرية للنموذج الملاحظ، و N حجم العينة.

ويؤكد هيو وبلنتر (Hu & Bentler, 1998) على ان هذا المقياس لا ينصح باستخدامه كثيرا بسبب حساسيته الشديدة لحجم العينة، فحينما تزيد حجوم العينات عــن 200 فإنــه لا يعطي أي دلالة إحصائية على جودة المطابقة، ويوضح الجدول 1 ملخصاً لكافة مؤشــرات جودة المطابقة المستخدمة في الكشف عن أحادية البعد تبعا للاسم والرمز وعلامــة القطــع المتفق عليها في الأدب النظري بقبول النموذج من رفضه (Bentler & Bonett, 1980).

جدول1: مؤشرات جودة المطابقة المستخدمة في التحليل العاملي التوكيدي

اسم المؤشر	الرمز	علامة القطع للمطابقة الجيدة
مؤشر مربع كاي Relative χ^2 / df	χ^2 / df	2:5 غير مطابق أقل من 2 مطابق قريب من 1 مطابق تماما
مؤشر حسن المطابقة Goodness of Fit Indexes	GFI	أكبر من 0.95
مؤشر حسن المطابقة المعدل Adjusted Goodness of Fit Indexes	AGFI	أكبر من 0.95
مؤشر حسن المطابقة المتشدد Parsimony Goodness of Fit Indexes	PGFI	بين 0.90 و 0.95
مؤشر جذر متوسط مربع البواقي The Root-mean-square Residual	RMR	القيمة أقل من 0.05
مؤشر جذر متوسط مربع الخطأ التقريبي The root-mean-square error of approximation	RMSEA	بين 0.07 - 0.03 تطابق ممتاز أقل من 0.01 تطابق تام
مؤشر المطابقة المعياري Normed Fit Index	NFI	بين 0.90 و 0.95 مطابق أكبر من 0.95 مطابق بصورة كبيرة
مؤشر المطابقة النسبي Relative Fit Index	RFI	أكبر من 0.90
مؤشر المطابقة المقارن Comparative Fit Index	CFI	أكبر من 0.95 أو 0.9
مؤشر توكر لويس Tucker–Lewis index	TLI	أكبر من 0.95 أو 0.9
مؤشر المطابقة المتزايد Incremental Fit Index	IFI	أكبر من 0.95 أو 0.9
مؤشر هولتر لجودة المطابقة بمستوى 0.05 Hoelter's (1983)`critical N' for a significance level of .05	Hoelter's	أكبر من حجم العينة المحدد

وعادة يستخدم أسلوب الأرجحية العظمى في تقدير كافة المعاملات بين المتغيرات وكافة مؤشرات جودة المطابقة، وللحكم على جودة نموذج معين يجب أن يحقق النموذج المفترض للمقياس مؤشرات جودة المطابقة، ويمكن الحكم على صدق عباراته أو صدق أبعاده في ضوء الأوزان الانحدارية المعيارية والتي تعرف بمعاملات الصدق أو التشبع على المتغير الكامن، وهنا يمكن تحديد قيمة معينة يتم على أساسها رفض التشبعات التي تقل عنها، وكذلك يمكن الحكم على صدق العبارات في ضوء النسبة الحرجة (Timothy, 2006).

تستخدم المؤشرات السابقة (مؤشرات جودة المطابقة) لمقارنة النماذج البنائية معـــا من خلال برمجيات وحزم إحصائية معقدة، في حين إن توفر نموذج الباحـــث فقــط فــأن البرمجيات الإحصائية وبصورة افتراضية by default تقوم بمقارنة النموذج الذي تم بنــاءه من قبل الباحث مع النموذج المستقل أو ما يسمى بالنموذج الصفري وهو النمـــوذج الـــذي تكون فيه المتغيرات الكامنة غير مترابطة والذي توفره البرمجية ذاتهـــا (Hu & Bentler , 1998).

العوامل التي تؤثر في مؤشرات جودة المطابقة (Factors that Affect Fit Indices):

يؤكد كيني (Kenny, 2012) على وجود عدة عوامل تؤثر في جودة المطابقة ويمكن إجمالها فيما يلي:

1- عدد المتغيرات الداخلة في النموذج (Number of Variables): يرى كيني وماكوش (Kenny & McCoach, 2003) إن زيادة عدد المتغيرات الداخلة في بناء النموذج تؤثر في جودة المطابقة، حيث تزداد قيمة المؤشر RMSEA لدى زيادة عدد المتغيرات في النموذج في حين تبقى قيم المؤشرات TLI و CFI مستقرة نسبيا، ولكن تميل إلى الانخفاض قليلا، كما تنخفض قيم المؤشر AGFI بزيادتها، ولكن لا يزال الأمر غير مفهوم لماذا النمـــاذج تصبح ضعيفة الجودة (Poor Fit) بزيادة عدد المتغيرات.

2- درجة تعقيد النموذج (Model Complexity) :أشار كل من اجزان ومادن (Ajzen, & Madden, 1986) أن أي تغيير في درجات الحرية يؤدي إلى تغيير في قيم المؤشـــرات، وأن درجة تعقيد النموذج سببا في نقصان قيم مؤشرات جودة المطابقة.

3- حجم العينة (Sample Size) : يؤكد كيني وماكوش (Kenny & McCoach, 2003) أن زيادة حجم العينة تؤثر في جودة المطابقة، نماذج بحجوم عينات كبيرة تؤدي لقـــيم أقـــل لمؤشرات جودة المطابقة، حيث تصبح قيم المؤشرات (RMSEA , RMR) أقل بزيادة حجم

العينة، في حين أن قيم المؤشر (TLI) يتأثر قليلا بنقصان حجم العينة ولكنه يحافظ على قيمته نسبيا، ولا يتأثر المؤشر CFI بزيادة أو نقصان حجم العينة، ويرى تانكا Tanaka, (1987) أن حجم العينة (N > 200) هو الحد الأدنى لحالات جودة المطابقة، وزيادة حجم العينة يعني زيادة الخطأ من النوع الثاني، وهذا لا يحقق الهدف المطلوب.

4- التوزيع الطبيعي للبيانات (Normality): يرى ريس (Reis, 1986) أن البيانات غير الطبيعية وخاصية التفلطح الكبير يؤثر في زيادة كبيرة في قيم مربع كاي، كما وتصبح مؤشرات جودة المطابقة المتزايدة والمقارنة أقل تأثيرا لجودة المطابقة.

5- طول الاختبار (Test Length): حيث أكد كل من ماك دونلاند ومولاك (McDonald & Mulaik, 1979). أنه على الأقل لا بد من وجود 15 فقرة لقياس قدرة واحدة، وان الاختبار الطويل المتعدد الأبعاد لا بد أن يحتوي على ثلاث فقرات على الأقل لتعريف البعد الواحد. وأن بعض المؤشرات تتأثر بزيادة طول الاختبار، حيث تقل قيمة كل من المؤشر NFI و RFI و PCFI بزيادة طول الاختبار.

الدراسات السابقة

يعتبر موضوع افتراض أحادية البعد من المواضيع الهامة جدا وخاصة في مجال قياس السلوك الإنساني والنفسي وأثره واضح على العديد من الاختبارات والمقاييس فتعددت الدراسات حوله. وبعد التقصي والبحث – في حدود علم الباحثة – تبين انه لم يتم إجراء أي بحث عربي يهتم بفاعلية المؤشرات المستخدمة في الكشف عن أحادية البعد ضمن إجراءات التحليل العاملي التوكيدي، وعلى الرغم من محاولات الباحثة العثور على دراسات تتعلق بالقضية المطروحة بهذا البحث بشكل مباشر، لم تجد دراسة عربية واحدة تتصل بهذا الموضوع، وأغلب الدراسات الوارد ذكرها في هذا الجزء هي دراسات أجنبية تتناول بعض جوانب القضية المطروحة ولم تحصى كافة مؤشرات الكشف عن أحادية البعد بصورة واضحة، وكان من الضروري استعراض الدراسات والجهود السابقة التي بذلت في هذا الصدد.

لذا يتناول هذا الفصل الدراسات السابقة المنشورة والتي تم الرجوع إليها حول موضوع افتراض أحادية البعد. وقد تم تصنيفها حسب ارتباطها بموضوع الدراسة كما يلي:

دراسات توضح أثر انتهاك افتراض أحادية البعد على بعض البرامج التقديرية:

قام كيرسي وهسو (Kirisci & Hsu, 1995) بدراسة حول قوة برنامج (BILOG)، عند انتهاك افتراض أحادية البعد لفقرات الاختبار والتوزيع الطبيعي للقدرة، ولتحقيق ذلك تم استعمال النموذج اللوجستي الثلاثي المعلم المتعدد الأبعاد، والذي يعد تعميما مباشرا للنموذج اللوجستي ثلاثي المعلم أحادي البعد لتوليد البيانات. وتم تغيير العوامل من حيث: أبعاد الاختبار، شكل توزيع القدرة، ودرجة الارتباطات الداخلية بين القدرات (0.0) أو (0.60)، وافترض أن توزيع معلمة التمييز للفقرة يقع ضمن الفترة (0.4 – 2) ومعلمة الصعوبة (2- - 2)، ومعلمة التخمين (0.0 – 0.30)، وافترض ثلاث توزيعات للقدرة : الطبيعي (الالتواء = 0.0،

والتفلطح = 0.0)، وملتو التواء موجب (الالتواء = 0.75 ، والـــتفلطح = 0.0)، وتوزيـــع منبسط (الالتواء = 0.0 ، والتفلطح = 1-).

وتم توليد الاستجابات على الفقرات، وأنشئت المصفوفة الاستجابات برتبـــة (40 X 1000)، بحيث يشير العدد 40 إلى عدد الفقرات، والعـــدد 1000 إلـــى عـــدد المفحوصـــين، وضوعفت شروط التجربة عشر مرات إذ ولدت عشر عينات لكل مجموعة ممن يمتلكـــون خصائص التوزيعات المطلوبة، واستخدم برنامج BILOG لتقدير معـــالم الفقـــرة كمـــا تـــم الحصول على نتائج من برنامج ASCAL وبرنـــامج MULTILOG لأغـــراض المقارنـــة، واستخدم جذر متوسط مربع الأخطاء RMSE (Root Mean Square Errors) الذي يعتمـــد على الفروق بين المعالم الحقيقية والمقدرة وتحليل التباين الثلاثي لإظهار النتائج.

أظهرت النتائج المرتبطة بقوة برنامج BILOG عند انتهاك افتراض أحاديـــة البعـــد لفقرات الاختبار، والتوزيع الطبيعي للقدرة، أن بعدية الاختبار لها تأثير دال في تقدير معالم الفقرة والقدرة، إذ يقدم الاختبار أحادي البعد نتائج أفضل في تقدير معلمة التمييز ومعلمـــة الصعوبة للفقرة، حيث كان (RMSE) للاختبار أحادي البعد أقل دلالة في الاختبار ثلاثـــي الأبعاد عند مستوى الدلالة (α = 0.05) على العكس من الاختبار ثلاثي الأبعاد الذي أنـــتج تقديرات أفضل في تقديره لمعلمة التخمين للفقرة، وعندما وجد ارتباطات داخلية بين القدرات كان (RMSE) للاختبار الثلاثي الأبعاد أقل دلالة من الاختبار أحادي البعد عنـــد مـــستوى الدلالة (α = 0.05)، وينتج الاختبار أحادي البعد تقديرات أفـــضل فـــي معلمتـــي التمييـــز والصعوبة ومعلمة القدرة، عندما تكون القدرات غير مترابطة.

وكان برنامج BILOG أقل قوة عند انتهاك افتراض شكل توزيع القدرة، في تقــدير معلمة التخمين للفقرة، ومعلمة القدرة للاختبار أحادي البعد، أما شكل توزيع القدرة الملتــو التواءً موجباً أعطى تقديرات ضعيفة في تقدير معلمة التخمين للفقرة، وعندما تكون أبعــاد القدرات مترابطة والاختبار متعدد الأبعاد.

وقورنت تلك النتائج مع النتــائج المستخلــصة مــن برنــامج ASCAL وبرنــامج MULTILOG عن طريق تمثيل (RMSE) بيانيا لمعالم الفقرة والقدرة في جميع البــرامج المستخدمة، واظهر برنامج BILOG أقل قيمة لــ (RMSE) لمعالم القدرة عنــد إهمــال بعدية الاختبار. وقدمت البرامج الثلاثة نتائج متشابهة عندما يكون شكل توزيع القدرة ملتــو التواءً موجباً، وكان أداء برنامج BILOG أفضل من البرامج الأخرى في تقدير معالم الفقرة والقدرة عندما تكون القدرات غير مترابطة.

وفي دراسة أخرى أجراها كيرسي وهسو ويو (Kirisci , Hsu & Yu, 2001) هدفت لتقييم حساسية تقديرات المعالم أحادية البعــد، المــشتقة مــن برنــامج BILOG وبرنــامج MULTILOG وبرنامج XCALIBRE، عندما يتم انتهاك افتراض أحادية البعد، وعنــدما لا يكون شكل توزيع القدرة طبيعيا، وبالتحديد فحصت الدراسة العوامــل الآتيــة: اختبــارات أحادية البعد، واختبارات ثلاثية الأبعاد بأبعاد مترابطة، شكل توزيع القدرة: طبيعي، ملتــو التواء موجبا وسالبا، برنامج التحليل المستخدم، وتم توليد البيانات لــ (1000 مفحوص في اختبار مكون من 40 فقرة) باستخدام النموذج اللوجستي الثلاثي المعلم المتعدد الأبعاد، وباستخدام البرامج الثلاثة السابقة تم تقدير معالم الفقرة والقدرة، واستخدم لأغراض التقييم الجذر التربيعي لمتوســط

مربعات الأخطاء للفروق بين قيم المعالم المقدرة والمعالم الحقيقية، نظرا لشيوع استعماله في دراسة تحسن المعالم.

أظهرت النتائج وجود تفاعل بين البرامج المختلفة وبعدية الاختبار، وكانت بـــرامج التقدير أحادية البعد أكثر حساسية لاختلاف شكل التوزيع وكانت مربعات الأخطاء لتقـــدير المعالم أقل ما يمكن في برنامج BILOG مقارنة بـــالبرامج الأخـــرى، وأظهـــر برنــــامج MULTILOG وبرنامج XCALIBRE تباينا قليلا في تقدير المعالم بسبب انتهاك افتـــراض أحادية البعد، باستثناء تقدير معلمة التمييز من خلال برنامج MULTILOG، وبالرغم مـــن اعتماد البرامج الثلاثة على الأرجحية القصوى في تقـــدير المعـــالم إلا أن نتائجهـــا كانـــت متناقضة.

دراسات توضح الطرق الاختبارية للتحقق من أحادية البعد:

ومن الدراسات الرائدة للمقارنة بين أساليب قياس افتراض أحادية البعد تلك الدراسة التي أجراها هامبلتون ورفينلي (Hambleton & Rovinelli,1986) لتقييم افتـــراض أحاديـــة البعد لمجموعة من فقرات الاختبار والتي تناولت المقارنة بين أربع طرق في الكشف عـــن افتراض أحادية البعد وهي: التحليل العاملي الخطي، تحليـــل البـــواقي، التحليـــل العـــاملي اللاخطي، وطريقة بيجار، ولتحقيق ذلك تم توليد بيانات من خمس مجموعات كل منها يتكون من 40 فقرة و 1500 مفحوص وحسب النموذج اللوجستي الثلاثي المعلم. وتم اعتماد معامل الارتباط بين السمات (0.1 أو 0.6) كما استخدمت المحكات السابقة في الكشف عن افتراض أحادية البعد.

وتم افتراض أحادية البعد في مجموعة البيانات الأولى، بمعامل ارتبـــاط 0.1 بـــين مجموعة البيانات الثانية والثالثة، ومعامل ارتباط 0.06 بين مجموعة البيانات الرابعة والخامسة،

وكانت أول 20 فقرة في مجموعة البيانات الثانية والرابعة تقيس سمة واحدة وبقية الفقــرات تقيس سمتين، أما مجموعة البيانات الثالثة والخامسة كانت أول 30 فقرة تقيس سمة واحــدة وبقية الفقرات تقيس سمتين. وحددت قيم معالم الصعوبة من 2 إلى 2-، وقيم معالم التمييــز من 2 إلى 0.4، ومعلمة التخمين كانت 0.25.

وقد توصلت الطرق الأربعة لنتائج متناقضة عن افتراض أحادية البعد، فقد أظهــر التحليل العاملي الخطي ضعفا في تقدير أحادية البعد في البيانات، أما طريقتي تحليل البواقي وبيجار لم تستطيعا الكشف عن افتراض أحادية البعد بنجاح خصوصا حينما يكون الارتباط بين السمتين ضعيفا، في حين كانت طريقة التحليل العاملي غير الخطي أكثر نجاحــا فــي الكشف عن افتراض أحادية البعد.

وفي دراسة أجراها لي (Lee, 2000) هدفت إلى تحديد مدى مناسبة تبني مفهــوم الاختبار الجزئي في تقدير الخطأ المعياري المشروط في قياسات نموذج المعادلة البنائيــة، لاختبارات مكونة من مجموعة اختبارات جزئية، وذلك بمقارنة الفروق في التقديرات بــين الطرق المبنية على الفقرة، وتلك المبنية على الاختبار الجزئي، وكشف الباحث عن افتراض أحادية البعد وافتراض الاستقلال الموضعي، واستخدم الباحث بيانات اختبار أياوا للمهارات الأساسية (Iowa Test of Basic Skills (ITBS لعام 1995، المكون من اختبار الاســتيعاب القرائي، واختبار الخرائط، والرسوم الإيضاحية لكل من الصفين الرابع والسابع، واختبــار المفردات للصف الثامن.

وتكونت عينة الدراسة من 985 طالبا من طلبة الصف الرابــع خــضعوا لاختبــار الاستيعاب القرائي، و914 من طلبة الصف الرابع خضعوا لاختبار الخرائط، و629 من طلبة

السابع خضعوا لاختبار الاستيعاب القرائي، و682 من طلبة الصف السابع خضعوا لاختبار الخرائط، و666 من طلبة الصف الثامن خضعوا لاختبار المفردات، واستخدم الباحث التحليل العاملي ممثلا بتحليل المكونات الرئيسية في الكشف عن أحادية البعد، واستخدم الإحــصائي Q3 وهو معامل ارتباط البواقي للكشف عن فرض الاستقلالية الموضعية، وأظهرت النتائج انه يتم انتهاك كل من افتراض أحادية البعد وافتراض الاستقلال الموضعي، عند التعامل مع الفقرة كوحدة أساسية في التحليل في الاختبارات المكونة من اختبارات جزئية.

وقد أجرى الشافعي (2008) دراسة حول تأثير انتهاك افتراضي أحاديـة البعـد واستقلالية الموضع في تدرج بنك الأسئلة ودقة معادلة درجات الاختبارات البنكية المسحوبة. تكونت عينة الدراسة من (505) مفحوصا، صمم الباحث من خلالها اختبـارين للدراسـة، يحتوي كل منها على 60 فقرة حول مواضيع الجبر والهندسة الفراغية من نوع الاختيار من متعدد، كان الاختبار الأول يشمل فقرات من مواضيع الجبر (ذات بعد واحد)، فـي حـين يحتوي الاختبار الثاني خليط من الجبر والهندسة (متعددة الأبعاد) وكان من نوع الاختيـار من متعدد والصواب والخطأ وإكمال الفراغ، وقسم الاختبارات إلى اختبارات فرعية بحيـث جعل نصفها متساوية في الصعوبة لأغراض المعادلة الأفقية، والأخرى مختلفة الـصعوبة لأغراض المعادلة العمودية، وقد استخدم الباحث المؤشرات التالية الخطأ المعياري للقياس، التحليل العاملي، تحليل القياسات المتكررة؛ وذلك للكشف عـن افتـراض أحاديـة البعـد واستقلالية الموضع، أظهرت النتائج وجود فروق ذات دلالة إحصائية لصالح الاختبار الأول والذي يضم فقرات ذات بعد واحد، وان انتهاك افتراض أحادية البعد قد أثر على دقة نتـائج المعادلة ودقة تدرج الفقرات للاختبار الثاني والذي يحتوي على مزيج من الجبر والهندسة، في حين لم تظهر أي دلالة إحصائية لانتهاك افتراض الاستقلال الموضعي.

في دراسة أجراها مكجيل (McGill, 2009) حول الطرق الاختبارية للتحقــق مــن أحادية البعد باستخدام التحليل للمركبات الرئيسية في ظل نظرية السمات الكامنة. اعتمد فيها الباحث على البيانات المولدة، واستخدم المؤشرات الآتية للتحقق من افتراض أحادية البعد: مؤشر كيزار the Kaiser (k>1) Criterion، النسبة المئوية لنــسبة التبــاين (greater than 50%, 40%, or 20%)، نسبة الجذور الكامنة، طريقة كلي Kelley method، بالإضافة إلــى مؤشر جديد استحدثه الباحث نفسه وأسماه باسمه.

وقد درس الباحث عدداً من العوامل وهي طول الاختبار: (20, 40, 60) فقرة، وحجم العينة (100, 250, 500, 750, 1000) مفحوصا، عندما توزعت القدرة بشكل طبيعي. وأثر ذلك على حساب كل من الخطأ من النوع الأول والخطأ من النوع الثاني وقوة الاختبار، وأثــره على التحقق من افتراض أحادية البعد. توصل الباحث إلى أن مؤشر كيزر المعدل كان أكثر فاعلية بزيادة عدد المفحوصين في الكشف عن افتراض أحادية البعد وكان حساب الخطأ من النوع الأول يقل بزيادة أعداد المفحوصين بهذه الطريقة وله علاقة بقوة الاختبار، وقد كــان ضعيف الكشف بنقصان أعداد المفحوصين. وتلتها قوة طريقة الجذور الكامنة في الكــشف عن افتراض أحادية البعد وذلك بزيادة عدد المفحوصين، ولم يزود أثر تغيير حجم الاختبار أي دلالة إحصائية حول افتراض أحادية البعد.

وفي دراسة لسوزان وسلوكم و برونو وزومبــو (Suzanne, Slocum, Bruno & Zumbo. 2010). حول تقييم افتراض أحادية البعد في المقاييس النفسية باســتخدام معــايير متعددة في التحليل العاملي، واعتمد الباحثون فيها على المؤشرات الآتية: مربــع كــاي χ^2، الجذر التربيعي لمتوسط الخطأ التقريبي، والجذور الكامنة، والتحليل المتوازي، ونسبة تباين الشيوع. وتم الاعتماد على البيانات المولدة بالاعتماد على نظرية استجابة الفقرة تحت ظروف

مختلفة من حجم العينة وتوزيع البيانات، أظهرت النتائج انه لا يوجد طريقة فاعلة للمفاضلة بينها جميعها في الكشف عن افتراض أحادية البعد، إلا أنها أبدت ضعفا واضحا في حالة الحجوم الصغيرة، وكان تباين الشيوع متدنياً في حالة الالتواء السالب لقدرات المفحوصين.

دراسات توضح العوامل المؤثرة على افتراض أحادية البعد:

وأجرى ليو (Liu, 1992) دراسة لمعرفة أثر بعدية بيانات اختبار المولدة من نماذج الاستجابة للفقرة ثنائية البعد التعويضية، وغير التعويضية، على مطابقة البيانات للنموذج. وقد استخدم نموذج استجابة الفقرة الثلاثي المعلمة، ونموذج استجابة الفقرة ثنائي البعد التعويضي، ونموذج استجابة الفقرة ثنائي البعد غير التعويضي. وتم استخدام بيانات اختبار متعددة الأبعاد في نموذج أحادي البعد، كمحك لفحص انتهاك افتراض أحادية البعد. وحسبت عدد الفقرات غير المطابقة للنموذج أحادي البعد، كما حسب متوسط قيم مربع كاي لبيانات الفقرة أحادية البعد، كمؤشر على مطابقة البيانات للنموذج.

أظهرت النتائج أن البيانات المولدة من النموذج غير التعويضي تميل أكثر لأن تكون ثنائية البعد، وغير مطابقة للنموذج أحادي البعد ثلاثي المعلم، وأن البيانات المولدة من النموذج التعويضي تميل لأن تكون أحادية البعد، وغير مطابقة للنموذج أحادي البعد الثلاثي المعلم، وأن الارتباط بين السمتين الكامنتين، لا يؤثر على افتراض أحادية البعد، وليس له أثر على مطابقة البيانات للنموذج ، بالرغم من وجود فروق في أحادية البعد بين النموذج التعويضي والنموذج غير التعويضي.

في دراسة لكارون ومايكل وكالن (Karon, Michael & Kallen, 2009) بعنوان دراسة مطابقة: أثر عدد الفقرات وتوزيع البيانات على المحكات التقليدية في تقييم أحادية البعد لنماذج

استجابة الفقرة. اعتمد الباحث في دراسته على البيانات المولدة باستخدام برنامج WinGen2 حيث تم توليد اختبارين وبواقع (47 ، 10) فقرة، مقابل البيانات الحقيقيـة المـسجلة مـن الاختبـارات (PROMIS) Patient-Reported Outcomes Measurement Information System وبواقع 47 فقرة تم سحبها عشوائيا من بنك الاختبارات المخصص لـذلك، ومـن الاختبارات التيكفية المحوسبة CAT Computer Adaptive Testing وبواقـع 10 فقـرات. وكانت أعداد المفحوصين 745 لكل اختبار، وأعتمد الباحثون في دراستهم على مؤشـرات التحليل العاملي التوكيدي، وتم استخدم مؤشر جذر الخطأ التقريبـي (RMSEA)، ومؤشـر توكر لويس (TLI)،مؤشر المطابقة المتزايد (IFI)، ومؤشـر المطابقـة المعيـاري (NFI)، ومؤشر المطابقة المقارن (CFI)، للكشف عن افتراض أحادية البعد.

كانت نتائج البيانات المولدة أكثر فاعلية لدى جميع المؤشـرات المـستخدمة فـي الكشف عن افتراض أحادية البعد مقابل البيانات الحقيقية، ولدى مقارنة فاعلية المؤشـرات معا أظهرت النتائج فاعلية مؤشر توكر لويس ومؤشر المطابقة المعياري في حالة الحجـوم الكبيرة للفقرات (47) فقرة، والتوزيع الملتوي لقدرات الأفراد، في حين بينـت أن مؤشـر الجذر الخطأ التقريبي كان ضعيفا حيال الكشف عن أحادية البعد. كما أظهر مؤشر المطابقة المقارن تحسنا بسيطا حيال أعداد الفقرات المنخفضة، إلا انه لم يكشف عـن اثـر حجـم الفقرات على اثر تعدد الأبعاد، ولم تبدي الدراسة أي اهتمـام لأثـر النمـوذج اللوجـستي المستخدم.

وفي دراسة أجراها وانج ووانج (Wang & Wang, 2003) هدفت للتعرف على أثر تقدير معالم أحادية البعد المشتقة من بيانات ثنائية البعد، و ثلاثية البعد، على تقديرات معالم نموذج راش الأحادي المتعدد الأوجه، وعالجت الدراسة أربعة عوامل وهي: أبعاد الـسمة

(أحادي البعد ، ثنائي البعد، ثلاثي البعد)، حجم عينة المفحوصين المستخدمة (500 1000، ،
2000)، درجة الارتباط بين أبعاد السمة كانت علـــى النحـــو (0.0، 0.3، 0.7)، وأعتمـــد الباحث على مؤشر جذر الخطأ التقريبي (RMSEA) للكشف عن افتراض أحادية البعد.

أظهرت نتائج الدراسة أن عدم تحقق افتراض أحادية البعد يؤثر على دقـــة تقـــدير المعالم، وتوصلت النتائج لوجود أثر قليل لحجم العينة على جذر متوسط مربع الأخطاء، ولم تظهر أي دلالة لأبعاد السمة (أحادي البعد ، ثنائي البعد، ثلاثي البعد) على افتراض أحادية البعد.

التعقيب على الدراسات السابقة:

يتبين من الدراسات السابقة أن الاهتمام بافتراض أحادية البعد دليل على أهمية هذا الافتراض لما له من أثر كبير في تطوير المقاييس والاختبارات، ويلاحظ أن انتهاك افتراض أحادية البعد أو التغاضي عنه يؤثر على الخصائص الإحصائية للفقرات التي يجرى تحليلهـــا باستخدام النماذج الرياضية المختلفة التي ترتبط بنظرية استجابة الفقرة حيث أظهرت دراســـة كل من كيرسي وهسو ويو (Kirisci , Hsu & Yu, 2001) ودراسة الشافعي (2008) ودراسة كل من كيرسي وهسو (Kirisci & Hsu, 1995) إلى أن انتهاك افتراض أحادية البعد أثر على تقديرات برنامج التحليل للفقرات والأفراد. ويتضح من خلال نتائج أغلب الدراسات الـــسابقة كما هو الحال في دراسة وانج ووانج (Wang & Wang, 2003) التي أشارت إلى أن عدم تحقق افتراض أحادية البعد يؤثر في تقدير المعالم، و يؤثر بشكل أو بآخر على الخـــصائص الإحصائية والتقديرات الخاصة بفقرات التحليل، مما قد يؤثر بدوره على التدريج العام للاختبار.

كما يتضح من خلال الدراسات السابقة اعتمادها على عدد قليل من المؤشرات في تحققها من افتراض أحادية البعد، كما هو الحال في دراسة كل من مكجيل (McGill,2009)، لسوزان وسلوكم و برونو وزومبو (Suzanne L. Slocum-Gori • Bruno D. Zumbo,2010)، ودراسة كل من هامبلتون وروفينيل (Hambleton & Rovinelli,1986)، التي اعتمدت بشكل رئيسي على التحليل العاملي ومربع كاي والجذر التربيعي لمربع الخطاء التقريبي، ودراسة واحدة فقط وهي لكارون ومايكل وكالن (Karon , Michael & Kallen, 2009) استخدمت عدد بسيط من مؤشرات التحليل العاملي التوكيدي مع مقارنة لفاعلية تلك المؤشرات فيما بينها، ولم تكشف أغلب الدراسات السابقة عن فاعلية تلك المؤشرات المستخدمة للتحقق من افتراض أحادية البعد.

أما بخصوص الدراسات التي تناولت العوامل المؤثرة في افتراض أحادية البعد فقد كانت قليلة ولم تأخذ كافة العوامل بعين الاعتبار، كما هو الحال في دراسة وانج ووانج (Wang & Wang, 2003)، ودراسة كارون ومايكل وكالن (Karon , Michael & Kallen,2009) ودراسة مكجيل (McGill, 2009) التي اهتمت بأثر حجم العينة ونوع النموذج على افتراض أحادية البعد، ولم تتحقق من فاعلية المؤشرات المستخدمة في الكشف عن افتراض أحادية البعد.

لذا جاءت هذه الدراسة بصورة مختلفة لتلقي الضوء على فاعلية المؤشرات المستخدمة في الكشف عن افتراض أحادية البعد من خلال معرفة اثر طول الاختبار وشكل توزيع القدرة من بيانات مولدة للنماذج اللوجستية المختلفة. وخاصة أن معظم الدراسات ركزت في إجراءاتها على محكات تقليدية، ولم تأتي بالذكر على الطرق المعتمدة على التحليل العالمي ألتوكيدي، كما أنها لم تكشف بصورة مباشرة عن أثر طول الاختبار وشكل توزيع القدرة ضمن نماذج

استجابة الفقرة المطابقة للبيانات، وأهمية ذلك على افتراض أحادية البعد، علاوة على ذلـــك لا توجد دراسة عربية – في حدود علم الباحثة – تكشف عن افتراض أحادي البعد بأسلوب التحليل العاملي التوكيدي وتكشف عن فاعلية مؤشرات الكشف عن افتراض أحادية البعد.

الفصل الثالث

الطريقة والإجراءات

تناول هذا الفصل وصفا للأدوات والبرامج المستخدمة ووصفا لطريقة اختيار خصائص معلم القدرة ومعالم الفقرات وتوليد البيانات، كما تضمن وصفا للطرق والأساليب والمعالجات الإحصائية التي تم استخدامها في هذه الدراسة.

تم الاعتماد على البيانات المولّدة تبعاً لنماذج استجابة الفقرة (أحادي المعلم، ثنائي المعلم، ثلاثي المعلم)؛ لما توفره البيانات المولدة من ظروف معيارية يصعب الحصول عليها من البيانات الحقيقية، وتوزيع مناسب لقدرات المفحوصين، وتوزيعات مناسبة لمعالم الفقرات المستخدمة؛ أي أن البيانات المولدة تساعد في السيطرة على الظروف التي يتم فيها إجراء اختبارات الفاعلية، وهذا يتيح الفرصة لأن تكون قدرات المفحوصين موزعة بشكل مناسب عند جميع مستويات القدرة، واستخدام بنك من الفقرات بمواصفات معينة، بحيث تكون الصعوبة للفقرات موزعة بشكل مناسب على متصل الصعوبة؛ وذلك لتوفير فقرات عند كل مستوى من مستويات القدرة، حيث يصعب الحصول على هذه الظروف المعيارية في البيانات الحقيقة.

وتم الاعتماد على التحليل العاملي التوكيدي للدارسة كأحد أساليب نموذج المعادلة البنائية الذي ينجم عنه عدد من المؤشرات (مؤشرات جودة المطابقة)، وقد تمت مقارنة فاعلية هذه المؤشرات في ضوء علامات قطع خاصة بها متفق عليها في الأدب النظري، للكشف عن أحادية البعد.

أولا: وصف متغيرات الدراسة:

المتغيرات المستقلة: ويمكن توضيحها كما يأتي:

- طول الاختبار (60، 30، 15) فقرة.

- شكل توزيع القدرة (طبيعي، ملتو التواءً موجباً، ملتو التواءً سالباً).

- النموذج اللوجستي المستخدم (أحادي المعلم، ثنائي المعلم، ثلاثي المعلم).

المتغير التابع: وهو المؤشر الإحصائي المستخدم في الكشف عن افتراض أحادية البعد، وقد تم الإشارة إليها مسبقا في الأدب النظري، وهـــو: (MacCallum, Browne and Sugawara :1996, Kline, 2005)

1. مؤشر نسبة مربع كاي الى درجات الحرية Relative χ^2 / df
2. مؤشر حسن المطابقة GFI Goodness of Fit Indexes
3. مؤشر حسن المطابقة المعدل AGFI Adjusted Goodness of Fit Indexes
4. مؤشر حسن المطابقة المتشدد PGFI Parsimony Goodness of Fit Indexes
5. مؤشر جذر متوسط مربع البواقي RMR The Root-mean-square Residual
6. مؤشر الجذر التربيعي لمتوسط مربع الخطأ التقريبي RMSEA
7. مؤشر المطابقة المعياري NFI Normed Fit Index
8. مؤشر المطابقة النسبي RFI Relative Fit Index
9. مؤشر المطابقة المقارن CFI Comparative Fit Index
10. مؤشر توكر لويس TLI Tucker–Lewis index
11. مؤشر المطابقة المتزايد IFI Incremental Fit Index

12. مؤشر هولتر لجودة المطابقة بمستوى 0.05 Hoelter's (1983)`critical N' for a significance level of .05

الأدوات والبرامج المستخدمة في الدراسة:

برنامج WINGEN: وهو اختصار ل (Windows Software that Generates IRT Model Parameters and Item Responses).

وهو برنامج يولد بيانات لمعالم نماذج نظرية استجابة الفقـرة والاسـتجابات عـن الفقرات، ويستخدم هذا البرنامج لتوليد القيم الحقيقية لمعلـم القـدرة θ، ولمعـالم الفقـرات (التمييز، الصعوبة، والتخمين)، والذي يمكن من خلاله توليد بيانات أحادية البعد ومتعـددة الأبعاد، كما يمكن توليد معالم الفقرات بالاعتماد على معلم القدرة، ويمكن توليد معلم القدرة بالاعتماد على معالم الفقرات، ويتم اختيار عدد المفحوصين، وشكل التوزيع لمعلم القـدرة (طبيعي، بيتا، متماثل)، وتحديد عدد الفقرات اللازمة وعدد الاسـتجابات ونـوع النمـوذج (أحادي، ثنائي، ثلاثي) وتحديد شكل التوزيع لمعالم الفقرات، وتوليد الاسـتجابات اعتمـادا على معلم القدرة ومعالم الفقرات، ويمكن تخزين ملفات تحتوي على قيم القدرات، ومعـالم الفقرات، واستجابات الطلاب، ومعاملات ارتباط الفقرات بالاختبار.

برنامج SPSS:وهو اختصار ل Statistical Package for the Social Sciences

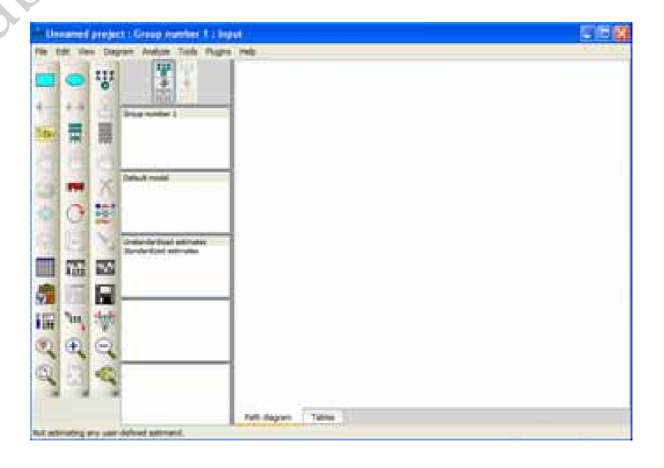
ويمثل الرزمة الإحصائية في العلوم الاجتماعية، وقد تم استخدام هذا البرنامج فـي الدراسة الحالية لحساب الإحصائيات الوصفية (الوسط الحسابي، الانحـراف المعيـاري، الالتواء، التفلطح).

برنامج AMOS 21: Analysis of Moment Structures

ويمثل برنامج لتحليل البناءات المقترحة آنياً، وتعد برمجية AMOS من إصدارات أل SPSS ويعتمد عليها معظم الباحثين في مجالات العلوم المختلفة وتخـدم هـذه التقنيـة الإحصائية نموذج المعادلة البنائية الذي يختبر صحة الفروض بين المتغيرات الكامنة (غير الملاحظة) والمتغيرات المقاسة (الملاحظة)، وهذا الأسلوب يستخدم في التحليل العـاملي التوكيدي للتحقق من الصدق البنائي للمقاييس المختلفة، ويتم من خلال هذه البرمجية بنـاء نموذج مقترح يربط بين تلك المتغيرات من خلال علاقات ارتباطيه، يتم دراستها واختبـار قوتها بناءً لنوع التحليل المطلوب.

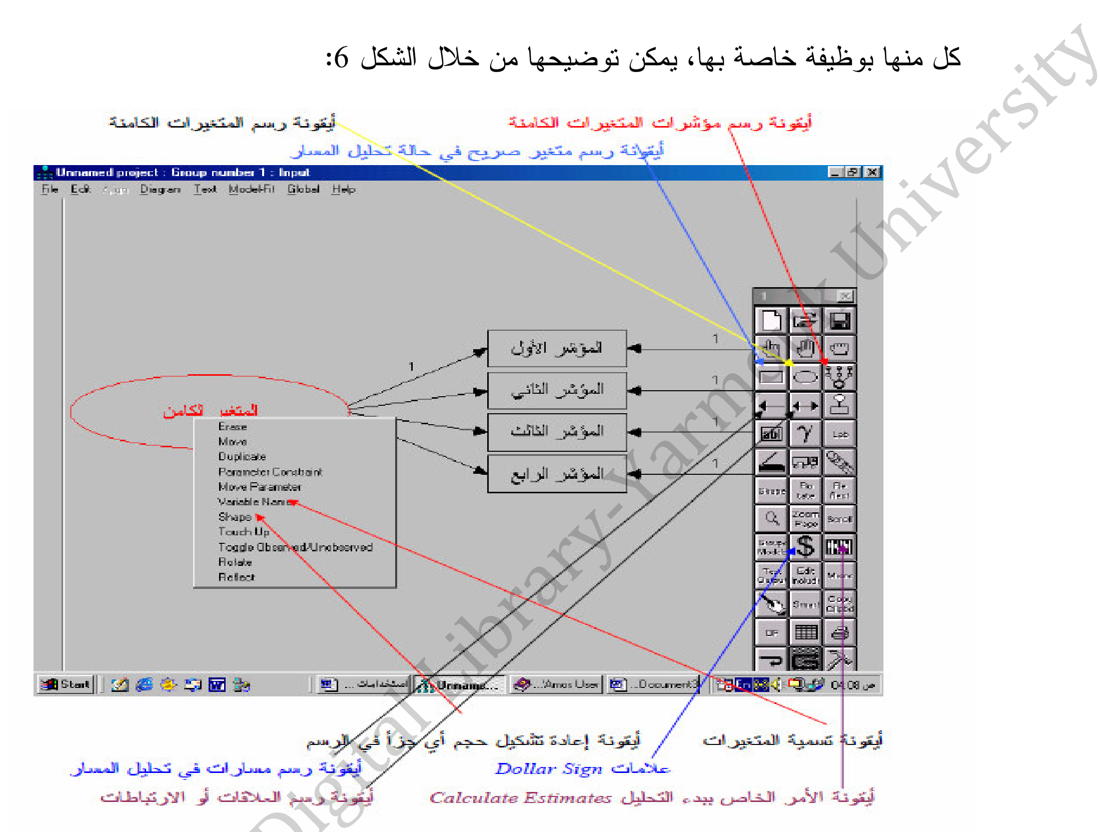
وهنا يظهر السؤال الذي يهتم التحليل التوكيدي بالإجابة عنه والمتمثل في: ما جودة النموذج المقترح (بين المتغيرات الكامنة والمتغيرات المقاسة)؟.

ويوضح الشكل 5 الواجهة الأساسية لبرمجية AMOS والتي تـساعد فـي بنـاء نموذج مقترح والعلاقات التي تربط بين المتغيرات الداخلة فيه، ويتبعه شرحا تفصيليا لكيفية بنائه:



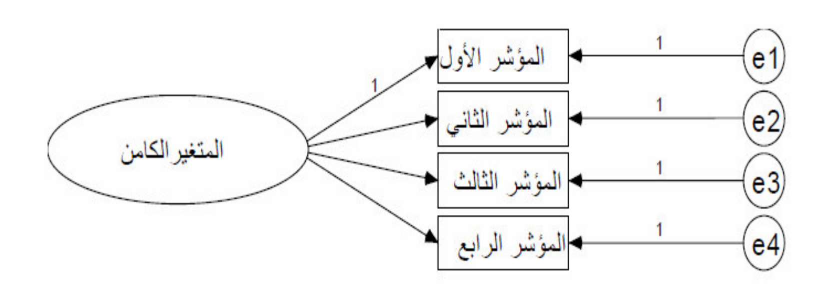
شكل 5 : رسم توضيحي للواجهة الأساسية لبرمجية AMOS

ويتم رسم النموذج المقترح الذي يربط المتغيرات معا من خلال الأيقونات، حيث تقوم كل منها بوظيفة خاصة بها، يمكن توضيحها من خلال الشكل 6:



شكل 6: واجهة برمجية AMOS مع توضيح عمل الأيقونات الأساسية

والشكل 7 يوضح شكل لنموذج مقترح بسيط تم رسمه باستخدام البرمجيـــة ذاتهـــا ويربط بين المتغير الكامن وعدد من المتغيرات الملاحظة، والعلاقات الارتباطية بينها متمثلة بأسهم.

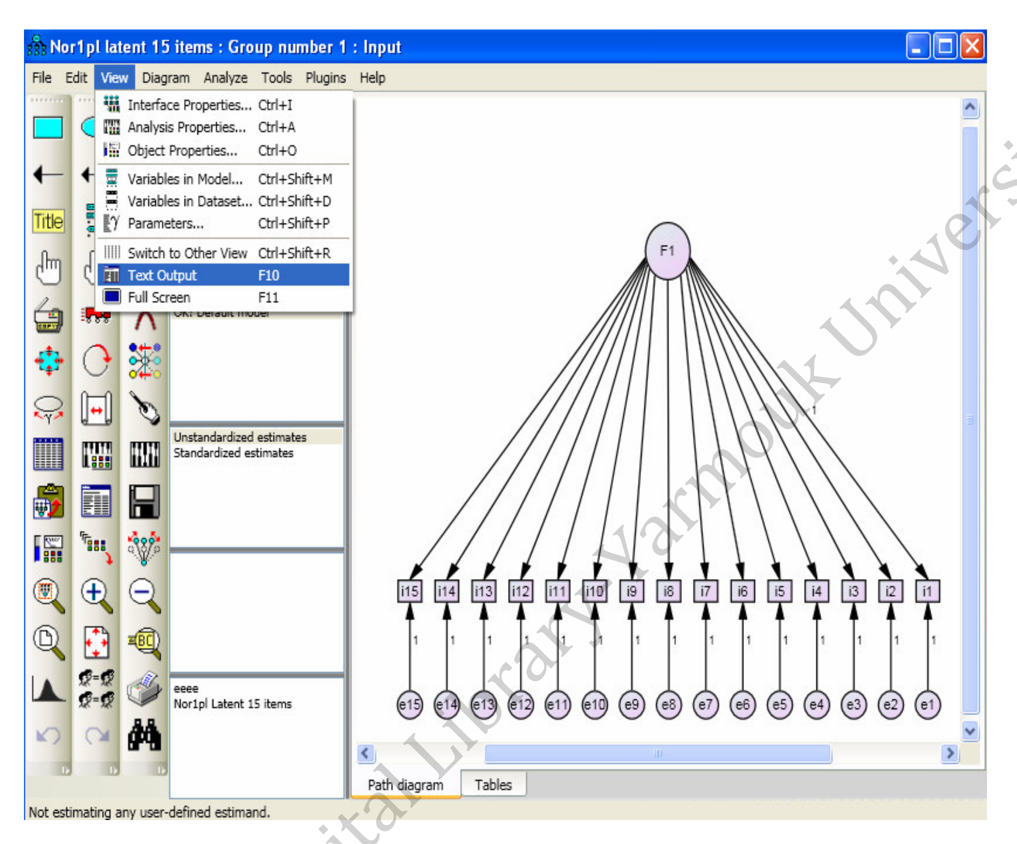


شكل 7 : رسم توضيحي لنموذج مقترح بسيط يربط المتغيرات معا

ويعبر عن المؤشر الكامن بالشكل البيضاوي، وتوجد في قائمة الأدوات الخاصة بالبرنامج أيقونة تحمل نفس الشكل وبمجرد الضغط على الماوس يتم رسم عدد من المتغيرات، أو فتح قائمة Diagram واختيار أمر draw unobserved variables وبالمثل يمكن رسم المتغيرات الملاحظة من خلال أمر draw observed variables، وتلك الأيقونات من أهم الأيقونات المستخدمة في الرسم، ويجب الانتباه إلى أن كل متغير يجب أن يحمل نفس الاسم الأصلي في ملف البيانات سواء من SPSS أو من Excel، ليتمكن البرنامج من قراءة البيانات، وبعد عملية الرسم واستدعاء البيانات يتم عمل تحليل إحصائي من خلال الأيقونة المشار إليها والخاصة ببدء التحليل Calculate Estimation.

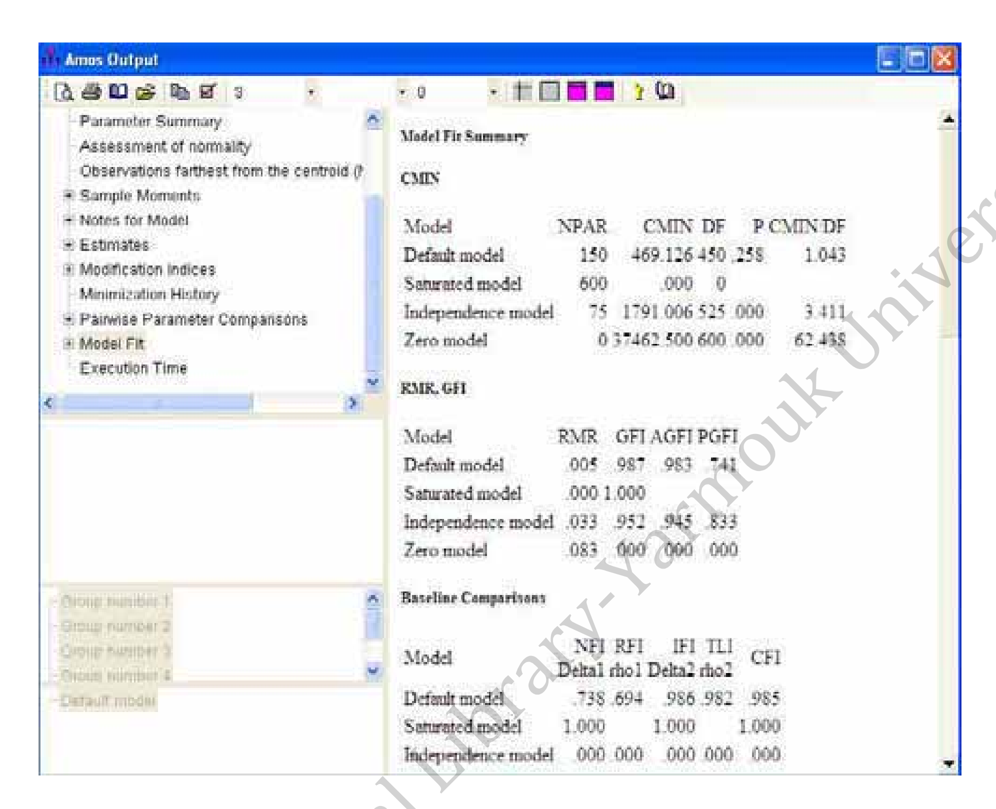
ويمكن من خلال البرمجية ذاتها بناء أكثر من نموذج مقترح يعتمد على عدد المتغيرات الكامنة وعدد المتغيرات المقاسة وطريقة ارتباطهما معا، وينتج العديد من المؤشرات على جودة المطابقة (بين مصفوفة التباين المصاحب للمتغيرات الداخلة في التحليل وبين المصفوفة المفترضة من قبل النموذج)، ويتم قبول النموذج أو رفضه في ضوء تلك المؤشرات.

وتقدم البرمجية السابقة عدداً كبيراً من مؤشرات جودة المطابقة تصل إلى 25 مؤشراً، يمكن استخرجها بعد الانتهاء من عملية التحليل من أيقونة view واختيار أمر text output، ويوضح الشكل 8 النموذج المقترح الذي تم بنائه وكيفية الوصول للمؤشرات:



شكل 8: رسم للنموذج المقترح الذي تم بناؤه فعلا لـ 15 فقرة توزعت بشكل طبيعي تبعا للنموذج أحادي المعلم

يبين الشكل 8 بناءً للنموذج المقترح والذي تم بناؤه فعليا والمتمثل في بناء نمـــوذج لـــ 15 فقرة توزعت فيها القدرة بشكل طبيعي تبعا للنموذج اللوجستي أحادي المعلمة، وبعد الضغط على الأيقونة view يتم الحصول على مخرجات عملية المطابقة من خلال اختيـــار أمر output text، والشكل 9 يوضح جزءاً من مخرجات مؤشرات جودة المطابقة الفعليـــة للنموذج المقترح مع توضيح كيفية قراءتها:



شكل 9 : رسم لمخرجات أو نتائج التحليل الإحصائي الفعلية / التحليل العاملي التوكيدي

وعادة تقدم برمجية AMOS ثلاثة مخرجات لنماذج مختلفة يمكن توضيحها كما يأتي (Kline , 1998):

1- النموذج المشبع (The Saturated Model).

2- النموذج المستقل The Independence Model.

3- النموذج الافتراضي The Default Model.

ويعتمد عادة على نتائج النماذج الافتراضية في تفسير نتائج مؤشرات جودة المطابقة للحكم على صحة النموذج. لذلك تم الاعتماد على النموذج الافتراضي للحكم علـى فاعليـة المؤشرات.

وقد تم رصد نتائج مصفوفة معاملات الارتباط، ومصفوفة التباينات المشتركة من خلال الضغط على أمر Estimates المشار إليه في الشكل السابق.

توليد البيانات:

تم توليد البيانات الخاصة باستجابة الأفراد باستخدام برمجية WINGEN، كما تم توليد مجموعة من الفقرات وفقاً لمتغيرات الدراسة، وكانت أعداد الأفراد (1000) فرداً، كما تم توليد تقديرات لقدرات الأفراد تبعا لمتغيرات الدراسة من حيث: طول الاختبار (15، 30، 60) فقرة، وتوزيعات شكل القدرة (طبيعية، ملتو التواءً موجبا (3.5 – 1)، وملتو التواءً سالبا (1 – 3.5))، بالاعتماد على النموذج اللوجستي (أحادي المعلمة وتراوحت فيه معلمة الصعوبة بين (2- - 2)، وثنائي المعلمة وتراوحت معلمة التمييز بين (4 - 0.5)، وثلاثي المعلمة وتراوحت معلمة التخمين بين (0.3 – 0.2))، وتم إخضاعها جميعها للتحليل الإحصائي.

القدرات:

تم استخدام برنامج WINGEN لتوليد تقديرات قدرات (1000) فرد، من توزيع طبيعي بوسط حسابي يساوي صفرا وانحراف معياري يساوي واحد، ولجميع متغيرات الدراسة؛ من حيث: طول الاختبار (15، 30، 60) فقرة، شكل توزيع القدرة (طبيعي، ملتو التواءً موجبا، وملتو التواءً سالبا)، ونوع النموذج اللوجستي المستخدم (أحادي المعلمة، ثنائي المعلمة، ثلاثي المعلمة)، تبعا للمواصفات السابقة، وتسمى هذه القدرات بالقدرات الحقيقية للأفراد، ويوضح الجدول الآتي الإحصائيات الوصفية لمعالم القدرة تبعاً لشكل التوزيع.

جدول 2: الإحصاءات الوصفية للقدرة الحقيقية

توزيع القدرة	القيمة الصغرى	القيمة العظمى	الوسط الحسابي	الانحراف المعياري	الالتواء	التفلطح
ملتو التواءً سالباً	-2.248	2.998	1.7459	1.0003	-1.016	.642
طبيعي	-3.531	3.63	-0.027	0.9885	-.131	.119
ملتو التواءً موجباً	-3	2.407	-1.659	1.0686	.977	.465

يلاحظ من الجدول 2 أن قيمة القدرة الحقيقية قد تراوحت بين 2.248- و 2.998 بوسط حسابي 1.7459 وانحراف معياري 1.0003 عندما توزعت القدرة بشكل ملتو التواءً سالبا، أما عندما توزعت القدرة بشكل طبيعي تراوحت قيمة القدرة الحقيقية بين 3.531- و 3.63 بوسط حسابي 0.027- وانحراف معياري 0.9885، في حين تراوحت قيمة القدرة الحقيقية بين 3.0- و 2.407 بوسط حسابي 1.659- وانحراف معياري 1.0686 عندما توزعت القدرة بشكل ملتو التواءً موجبا.

معالم الفقرات:

لتحقيق هدف الدراسة تم توليد مجموعة من الفقرات تبعا للمتغير المستقل طول الاختبار (15، 30، 60) فقرة، بالاعتماد على معلم القدرة الذي تم توليده مسبقا، وتبعا للنموذج اللوجستي (أحادي المعلم، ثنائي المعلم، ثلاثي المعلم)، وتمثل الجداول 5،4،3 ملخصا للإحصاءات الوصفية لمعالم الفقرات تبعاً لشكل التوزيع.

جدول 3: الإحصاءات الوصفية لمعلمة الصعوبة للنموذج أحادي المعلم

شكل توزيع القدرة	عدد الفقرات	القيمة الصغرى	القيمة العظمى	الوسط الحسابي	الانحراف المعياري	الالتواء	التفلطح
	15	-1.745	2.541	-0.002	1.2583	0.75745	-0.4043
ملتو التواءً سالباً	30	-2.233	2.393	0.1784	1.0075	-0.2373	0.1455
	60	-2.656	3.125	-0.032	1.1456	0.12392	0.0738
	15	-1.764	1.991	0.2771	1.0062	-0.2972	-0.0008
طبيعي	30	-1.542	1.701	0.1436	0.77	-0.3255	-0.1729
	60	-1.901	1.918	0.1154	0.9406	-0.1443	-0.4145
	15	-2.623	1.377	-0.395	1.2944	-0.3959	-0.7447
ملتو التواءً موجباً	30	-1.87	2.166	-0.135	1.0081	0.35742	-0.5243
	60	-1.946	2.176	0.2177	0.9956	-0.0799	-0.5966

يلاحظ من خلال الجدول 3 أن قيم معلمة الصعوبة تراوحت ضمن المدى المطلوب (2- ,2)، لمتغيرات الدراسة: طول الاختبار (15، 30، 60)، وشكل توزيع القدرة (ملتو التواءً سالبا، طبيعي، ملتو التواءً موجبا)، بوسط حسابي قريب من الصفر وبانحراف معياري قريب من الواحد صحيح، تبعا للنموذج أحادي المعلم، وهذا يحقق الغاية من الدراسة المطلوبة.

الجدول 4: الإحصاءات الوصفية لمعلمة الصعوبة والتمييز للنموذج ثنائي المعلم

شكل القدرة	عدد الفقرات	المعلمة	القيمة الصغرى	القيمة العظمى	المتوسط الحسابي	الانحراف المعياري	الالتواء	التفلطح
ملتو التواء سالب	15	b	-2.770	1.901	0.120	1.169	1.640	-0.725
		a	0.759	3.405	1.297	0.675	6.898	2.417
	30	b	-1.826	1.649	0.087	0.861	-0.320	0.031
		a	0.522	3.622	1.499	0.701	1.231	1.066
	60	b	-2.350	2.310	0.007	0.954	0.060	0.073
		a	0.581	3.386	1.614	0.657	0.225	0.893
طبيعي	15	b	-1.317	2.137	0.255	0.922	0.025	0.459
		a	0.734	3.524	1.433	0.644	8.680	2.647
	30	b	-1.154	2.225	0.209	0.861	-0.107	0.442
		a	0.423	3.315	1.578	0.824	-0.764	0.476
	60	b	-1.858	2.439	0.069	0.965	-0.108	0.283
		a	0.559	3.563	1.465	0.644	2.638	1.346
ملتو التواء موجب	15	b	-1.435	1.972	0.201	0.922	-0.143	-0.110
		a	1.114	4.229	2.231	1.006	-0.605	0.847
	30	b	-2.050	1.834	-0.045	0.796	0.794	-0.066
		a	0.730	4.159	1.769	0.766	1.762	1.090
	60	b	-2.035	2.092	0.029	0.954	-0.468	0.219
		a	0.493	5.471	1.806	0.886	3.547	1.439

حيث تشير b إلى معلمة الصعوبة، وتشير a إلى معلمة التمييز.

يلاحظ من خلال الجدول 4 أن قيم معلمة الصعوبة تراوحت ضمن المدى المطلوب (2- ,2)، لمتغيرات الدراسة: طول الاختبار (15، 30، 60)، وشكل توزيع القدرة (ملتــو التواءً سالبا، طبيعي، ملتو التواءً موجبا)، بوسط حسابي قريب مــن الــصفر وبــانحراف معياري قريب من الواحد صحيح، في حين أن قيم معلمة التمييز قد تراوحت ضمن المــدى (0.5, 5) ، ولجميع مستويات متغيرات الدراسة: طول الاختبار وشكل توزيع القدرة، بوسط حسابي تراوح بين (1.806 - 1.297)، وانحراف معياري (1.006 - 0.675)، تبعا للنموذج ثنائي المعلم، وهذا يحقق الغاية من الدراسة المطلوبة.

الجدول 5: الإحصاءات الوصفية لمعلمة الصعوبة والتمييز والتخمين للنموذج ثلاثي المعلم

شكل القدرة	عدد الفقرات	المعلمة	القيمة الصغرى	القيمة العظمى	المتوسط الحسابي	الانحراف المعياري	الالتواء	التفلطح
ملتوٍ التواء سالب	15	b	-1.903	2.169	-0.189	1.16	0.219	-0.401
		a	0.493	4.101	1.699	0.95	1.233	1.763
		c	0.211	0.298	0.257	0.03	-0.008	-1.483
	30	b	-2.261	1.904	-0.139	1.01	-0.112	-0.06
		a	0.453	3.123	1.57	0.76	0.456	-0.663
		c	0.203	0.296	0.248	0.03	0.012	-1.195
	60	b	-1.794	1.773	-0.014	0.95	0.006	-0.95
		a	0.495	6.06	1.742	1.01	1.764	4.714
		c	0.201	0.3	0.249	0.03	0.256	-0.809
طبيعي	15	b	-2.574	3.04	-0.269	1.44	0.695	0.857
		a	0.786	3.188	1.793	0.68	0.533	-0.354
		c	0.208	0.299	0.249	0.03	0.512	-0.399
	30	b	-1.983	1.389	-0.091	0.88	-0.391	-0.232
		a	0.633	4.27	1.584	0.85	1.367	1.863
		c	0.203	0.293	0.247	0.03	0	-1.106
	60	b	-1.975	2.021	-0.076	0.96	0.044	-0.429
		a	0.437	4.596	1.522	0.75	1.837	5.052
		c	0.201	0.297	0.25	0.03	0.055	-1.136
ملتوٍ التواء موجب	15	b	-2.207	0.921	-0.067	0.91	-1.077	0.662
		a	0.708	3.016	1.637	0.74	0.243	-1.12
		c	0.207	0.291	0.248	0.03	0.15	-1.088
	30	b	-2.775	2.732	0.199	1.31	-0.466	0.169
		a	0.415	4.347	1.807	0.92	1.302	2.073
		c	0.21	0.292	0.259	0.03	-0.757	-0.768
	60	b	-2.028	2.419	-0.022	0.89	0.329	0.725
		a	0.374	7.506	1.921	1.12	2.231	9.264
		c	0.2	0.3	0.251	0.03	0.173	-1.093

حيث تشير b إلى معلمة الصعوبة، وتشير a إلى معلمة التمييز، وتشير c إلى معلمة التخمين.

يلاحظ من الجدول 5 أن قيم معلمة الصعوبة تراوحت ضمن المدى (-2 ,2)، لجميع مستويات متغيرات الدراسة: طول الاختبار، وشكل توزيع القدرة، بوسط حسابي قريب من الصفر وبانحراف معياري قريب من الواحد صحيح، في حين أن قيم معلمة التمييز قد تراوحت ضمن المدى (0.5, 5)، ولجميع مستويات متغيرات الدراسة بوسط حسابي تراوح بين (1.522 –

1.921)، وانحراف معياري (1.12 – 0.68)، في حين تراوحت قيم معلمة التخمين بين (0.3 – 0.2)، بوسط حسابي يقع ضمن القيمتين السابقتين، بانحراف معياري ثابت 0.03 لجميع متغيرات الدراسة، تبعا للنموذج ثنائي المعلم، وهذا يحقق الغاية من الدراسة المطلوبة.

توليد الاستجابات:

بعدما تم توليد معلم القدرة كخطوة أولى، ومعالم الفقرات كخطوة ثانية، تم توليد استجابات الأفراد على الفقرات بناءً على معلم القدرة ومعالم الفقرات باستخدام برمجية WINGEN، وتم الاحتفاظ بها بملفات خاصة لاستدعائها وقت الحاجة.

ثالثا: بناء نموذج المعادلة البنائية

تم بناء نموذج المعادلة البنائية لكل متغير مستقل على حدا بمستوياته المختلفة، باستخدام برمجية AMOS 21، وبذلك تم بناء 27 نموذج مقترح لتغطية كافة جوانب مستويات متغيرات الدراسة الثلاث السابقة (طول الاختبار وشكل توزيع القدرة ونوع النموذج المستخدم(3 X 3 X 3))، كما تم رصد كافة النتائج التابعة لها (مصفوفة معاملات الارتباط، ومصفوفة التباينات المشتركة) ضمن جداول خاصة.

رابعا: التحليل العاملي التوكيدي

تم حساب كافة التحليلات الإحصائية الخاصة بالبيانات السابقة وإجراء التحليل العاملي التوكيدي لاستخراج مؤشرات جودة المطابقة باستخدام البرمجية السابقة، وبعد الحصول على كافة النتائج للمؤشرات المعتمدة تم تبويبها ضمن جداول خاصة لكل مؤشر على انفراد بحيث تمت المقارنة فيما بينها تبعا لعلامات القطع المحددة والمتفق عليها حسب الأدب النظري والبحوث السابقة.

واستخراج القيم التقديرية الإحصائية للمؤشرات المستخدمة في الكشف عن أحادية البعد. وتم تكرار عملية التحليل الإحصائي لاستجابات المفحوصين على الفقرات السابقة لخمس مرات متتالية لضمان ثبات القيم (Elias, 1997; Hattie, 1984, 1985, Kirisci & Hsu, 1995)، وللحصول على قيم للمؤشرات تتمتع بالثبات (Stability Replication)، ليتم اعتمادها بالشكل النهائي.

سجلت خمسة محاولات على اقل تقدير تبعاً لرأي تشرشل (Churchill, 1979)، لعملية التحليل الإحصائي لعينة الفقرات جميعها وتم رصد النتائج عليها لضمان ثبات التقديرات للقيم المعتمدة. وتم عرضها ضمن جداول خاصة فيها.

خامسا: مقارنة المؤشرات المستخدمة معا

تمت عملية مقارنة مؤشرات جودة المطابقة معا تبعا لكل مؤشر ولكل متغير مستقل على انفراد في الحالات جميعها، وتم عرضها في جداول خاصة للحكم على صدق النموذج المقترح كما سيأتي لاحقا.

سادسا: حساب الاختبار الإحصائي مربع كاي (Chi Square) لحسن المطابقة

تم حساب الاختبار الإحصائي مربع كاي لحسن المطابقة(χ^2) لكافة مؤشرات جودة المطابقة، من خلال تثبيت أثر متغيرين ودراسة اثر المتغير الثالث على مؤشرات جودة المطابقة، لتحديد ما إذا كانت الفروق بين التكرارات المتوقعة والملاحظة في أي عدد من متغيرات الدراسة السابقة تبعا لعلامات القطع المتفق عليها، ترجع منطقيا إلى اختلافات جوهرية بينها وذلك للحكم على فاعلية تلك المؤشرات لمطابقة البيانات والنموذج المقترح.

سابعا: المقارنة بين المؤشرات جميعها للحكم على جودة المطابقة وفاعلية المؤشرات.

تمت عملية المقارنة بين مؤشرات جودة المطابقة لكافة متغيرات الدراسة للحكم على مدى فاعليتها تبعا للاختبار الإحصائي مربع كاي لحسن المطابقة(χ^2)، وتم عرض ذلك في جداول إحصائية مبوبة.

ثامنا: تم الاعتماد على البرمجيات الإحصائية الآتية:

- SPSS
- WINGEN
- AMOS 21

في تقدير معالم الفقرة وقدرات المفحوصين، ولكافة العمليات الحسابية والتحليلات الإحصائية الخاصة بالدراسة. وتم عرض كافة النتائج وتبويبها ضمن جـداول إحـصائية خاصة ورصد كافة النتائج التي يتم الحصول عليها.

الفصل الرابع

النتائج

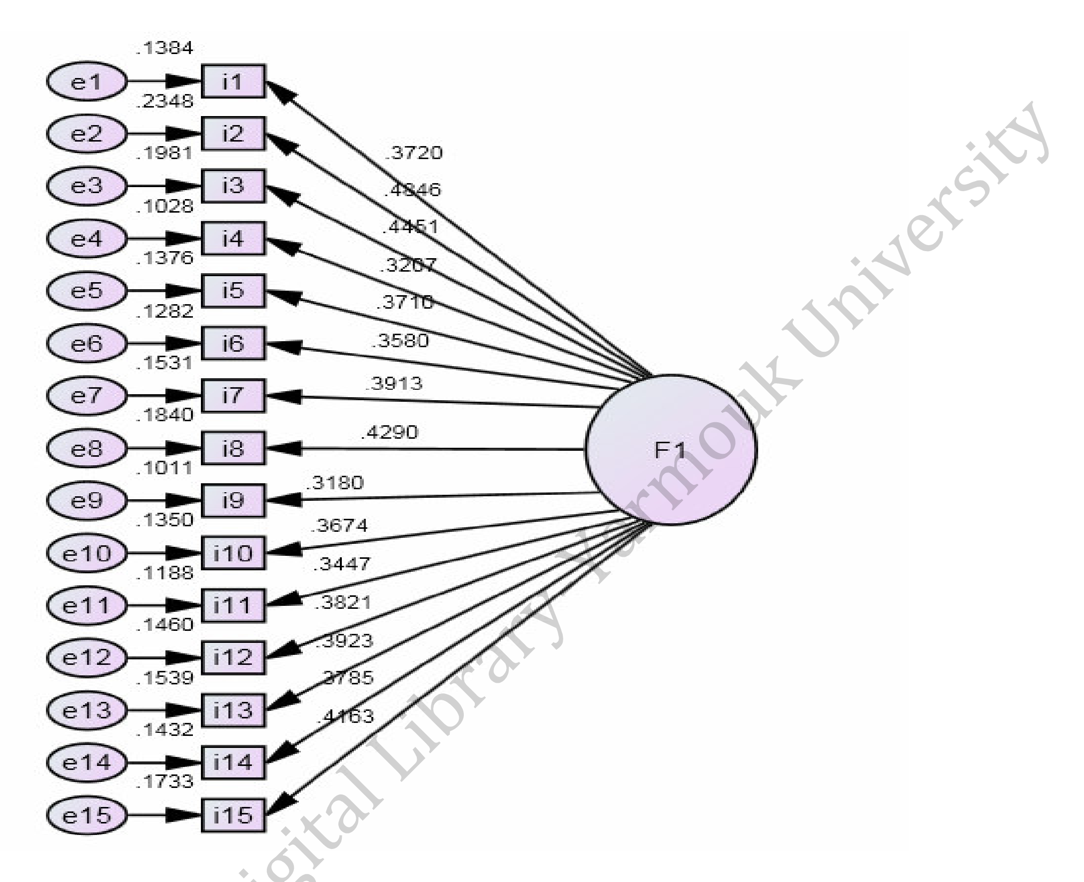
تناول هذا الفصل الوصف الإحصائي للنتائج التي توصلت إليها الدراسة، والتي هدفت للتحقق من فاعلية مؤشرات الكشف عن افتراض أحادية البعد وفق نماذج نظرية استجابة الفقرة، في ضوء تغير طول الاختبار وشكل توزيع القدرة والنموذج اللوجستي المستخدم. وتم الاعتماد على البيانات المولدة نظرا لحاجة الدراسة لأعداد كبيرة من المفحوصين. ولتحقيق هذا الهدف تم الحصول على البيانات اللازمة وإدخالها إلى ذاكرة الحاسوب، واسُتخدمت البرمجيات (SPSS, WINGEN, AMOS 21) لتوليد البيانات وتحليلها واستخراج النتائج، وفيما يلي عرض تفصيلي لهذه النتائج.

أولا: بناء نموذج المعادلة البنائية:

تم بناء نموذج المعادلة البنائية وفقا للمتغيرات المستقلة:

1. طول الاختبار (15، 30، 60) فقرة.
2. شكل توزيع القدرة (ملتو التواءً موجبا وملتو التواءً سالبا، طبيعي).
3. نماذج استجابة الفقرة (أحادي المعلم، ثنائي المعلم، ثلاثي المعلم).

وعلى هذا الأساس تم بناء 27 نموذجاً مقترحاً ودراستها بصورة منفصلة، والشكل الآتي مثالا توضيحيا لبناء نموذج هيكلي مقترح لبيانات تم توليدها وفقا للنموذج أحادي المعلم بطول اختبار 15 فقرة عندما توزعت القدرة بشكل طبيعي، علما أن بقية النماذج تم بناؤها بالطريقة نفسها باستخدام برمجية AMOS 21:



شكل10 : نموذج المعادلة البنائية المقترح لـ: 15 فقرة توزعت بشكل طبيعي وفقا للنموذج أحادي المعلم

حيث أن: $i_1, i_2, i_3, i_4, i_5, i_6, i_7, i_8, i_9, i_{10}, i_{11}, i_{12}, i_{13}, i_{14}, i_{15}$ تمثل طول الاختبار المكون من 15 فقرة وهي المتغيرات الملاحظة الداخلة في بناء النموذج.

كما تمثل e1, e2, e3, e4, e5, e6, e7, e8, e9, e10, e11, e12, e13, e14, e15 : الخطأ التابع لكل فقرة وهي المتغيرات غير الملاحظة (متغيرات البواقي : Residual Variables).

أما F1 تمثل السمة الكامنة (Latent variable) التي يتم دراسة التغير عليها وفقا لتلك المتغيرات، وهي هنا المؤشرات التي تستخدم للكشف عن افتراض أحادية البعد.

يلاحظ من الشكل 10 معاملات الارتباط بين السمة الكامنة والمتغيرات الداخليـــة، ومعاملات الارتباط بين المتغيرات الداخلية والمتغيرات غير الملاحظة(الخطأ).

جدول 6: مصفوفة الارتباطات للبيانات المولدة لـ : 15 فقرة توزعت بشكل طبيعي وفقا للنموذج أحادي المعلم

	i15	i14	i13	i12	i11	i10	i9	i8	i7	i6	i5	i4	i3	i2	i1
i15	1.00														
i14	.128	1.00													
i13	.195	.161	1.00												
i12	.184	.081	.137	1.00											
i11	.114	.130	.134	.130	1.00										
i10	.165	.127	.182	.134	.108	1.00									
i9	.113	.147	.092	.096	.106	.143	1.00								
i8	.186	.193	.152	.138	.114	.157	.146	1.00							
i7	.196	.145	.154	.172	.124	.172	.099	.190	1.00						
i6	.137	.143	.151	.130	.124	.133	.112	.134	.109	1.00					
i5	.129	.145	.132	.110	.073	.120	.121	.179	.149	.131	1.00				
i4	.149	.138	.071	.100	.159	.125	.049	.106	.113	.150	.134	1.00			
i3	.138	.172	.155	.181	.149	.114	.171	.205	.131	.160	.187	.153	1.00		
i2	.195	.163	.197	.205	.211	.170	.171	.154	.180	.143	.191	.100	.223	1.00	
i1	.140	.127	.117	.175	.128	.090	.094	.178	.109	.166	.138	.143	.171	.177	1.00

Eigenvalues (قيم الجذور الكامنة)

3.030 1.027 1.005 .964 .950 .898 .882 .866 .817 .813 .803 .788 .744 .733 .680

يبين الجدول 6 مصفوفة معاملات الارتباطات، حيث يمثل القطر الرئيسي مربع معامل الارتباط للمتغير نفسه ويساوي واحد صحيح، في حين تمثل الخلايا الأخرى معاملات الارتباط بين الفقرات المختلفة(المتغيرات المختلفة)، وقد تم حساب كافة الارتباطات بين المتغيرات السابقة وتحويلها إلى تقديرات معيارية وتم الحصول على مصفوفة التغاير

للمتغيرات، كما تم الحصول على كافة قيم الجذور الكامنة لجميـــع الفقـــرات. والجـــدول 7 يوضح مصفوفة التغاير للمتغيرات معا:

جدول 7: مصفوفة التغاير للبيانات المولدة لـ : 15 فقرة توزعت بشكل طبيعي وفقا للنموذج أحادي المعلم

	i15	i14	i13	i12	i11	i10	i9	i8	i7	i6	i5	i4	i3	i2	i1
i15	.244														
i14	.028	.201													
i13	.048	.036	.245												
i12	.041	.016	.030	.200											
i11	.025	.026	.029	.026	.195										
i10	.035	.024	.038	.025	.020	.180									
i9	.022	.026	.018	.017	.018	.024	.151								
i8	.046	.043	.038	.031	.025	.033	.028	.249							
i7	.048	.032	.038	.038	.027	.037	.019	.047	.249						
i6	.031	.029	.034	.026	.025	.026	.020	.030	.025	.206					
i5	.032	.032	.033	.025	.016	.025	.024	.045	.037	.030	.250				
i4	.027	.022	.013	.016	.026	.019	.007	.019	.020	.025	.024	.132			
i3	.034	.039	.038	.041	.033	.024	.033	.051	.033	.036	.047	.028	.250		
i2	.048	.036	.049	.046	.046	.036	.033	.038	.045	.032	.047	.018	.055	.247	
i1	.033	.027	.028	.038	.027	.018	.018	.043	.026	.036	.033	.025	.041	.043	.233

Eigenvalues (قيم الجذور الكامنة)

.677 .238 .225 .214 .210 .203 .195 .187 .185 .172 .166 .160 .156 .134 .112

يبين الجدول 7 مصفوفة التباين المصاحب (التغاير)، حيث يمثل قطر المصفوفة التباينات للمتغير نفسه، في حين تمثل قيم الخلايا الأخرى التباين المصاحب (التغاير) للمتغيرات الداخلة في النموذج، وقد تم الحصول على كافة قيم الجذور الكامنة لجميع الفقرات.

تم الحصول على مؤشرات جودة المطابقة من خلال إجراء التحليل العاملي التوكيدي باستخدام البرمجية ذاتها تبعا لمصفوفة التغاير، وكانت النتائج كما يلي:

ثانيا: النتائج المتعلقة بمؤشرات جودة المطابقة:

المؤشر الإحصائي مربع كاي (χ^2 / df):

لمعرفة قيم الإحصائي (χ^2) تم توليد بيانات بواقع (1000) مفحوص لكافة متغيرات الدراسة؛ طول الاختبار (15، 30، 60) فقرة، شكل توزيع القدرة (ملتو التواء موجب وملتو التواء سالب، طبيعي)، وفق نماذج استجابة الفقرة (أحادي المعلم، ثنائي المعلم، ثلاثي المعلم)، تم تحليل البيانات المولدة باستخدام برمجية AMOS 21 واستخراج مؤشرات جودة المطابقة، باستخدام التحليل العاملي التوكيدي للكشف عن أحادية البعد، تم التوصل لكافة النتائج المتعلقة بتلك المؤشرات.

ومن خلال دراسة المؤشر χ^2 لوحظ أنه عند مستوى الاحتمالية ($p > 0.05$) تكون الفروق غير دالة إحصائيا وهذا يعني تطابق البيانات المولدة للنموذج. وتجدر الإشارة بأن هذا المؤشر يتأثر بحجم العينة وبدرجات الحرية وقد يعطي تطابقا لدى العينات الكبيرة لذلك لابد أن يؤخذ في الاعتبار بعض المؤشرات الأخرى لجودة المطابقة بجانب النسبة بين قيمة χ^2 ودرجات الحرية df للكشف عن أحادية البعد (Hooper, Coughlan & Mullen, 2008). ويبين الجدول 8 عرضا تفصيليا للنتائج المؤشر χ^2/df :

جدول 8: قيمة المؤشر (χ^2 / df) للبيانات المولدة وفقا لنماذج استجابة الفقرة باختلاف طول الاختبار وشكل توزيع القدرة

النموذج	توزيع القدرة	عدد الفقرات	χ2	درجات الحرية	مستوى الاحتمالية	χ2/df
أحادي المعلم	ملتو التواء سالب	15	561.46	450	0.000	1.248
		30	2355.161	2025	0.000	1.163
		60	10531.641	8550	0.000	1.232
	طبيعي	15	469.126	450	0.258	1.043
		30	1995.421	2025	0.676	0.985
		60	8646.175	8550	0.230	1.011
	ملتو التواء موجب	15	458.709	450	0.378	1.019
		30	2381.12	2025	0.000	1.176
		60	11144.457	8550	0.000	1.303
ثنائي المعلم	ملتو التواء سالب	15	906.73	450	0.000	2.015
		30	3364.94	2025	0.000	1.662
		60	14890.275	8550	0.000	1.742
	طبيعي	15	854.986	450	0.000	1.900
		30	2846.737	2025	0.000	1.406
		60	9549.039	8550	0.000	1.117
	ملتو التواء موجب	15	1977.533	450	0.000	4.395
		30	3907.816	2025	0.000	1.930
		60	16849.446	8550	0.000	1.971
ثلاثي المعلم	ملتو التواء سالب	15	1239.14	450	0.000	2.754
		30	4329.469	2025	0.000	2.138
		60	18411.822	8550	0.000	2.153
	طبيعي	15	605.073	450	0.000	1.345
		30	2346.958	2025	0.000	1.159
		60	9870.82	8550	0.000	1.154
	ملتو التواء موجب	15	513.521	450	0.020	1.141
		30	2108.972	2025	0.095	1.041
		60	8892.062	8550	0.005	1.040

يتضح من الجدول 8 أن قيم مؤشر (χ^2) تشير إلى تطابق البيانات للنموذج عنـــدما توزعت القدرة بشكل طبيعي، على اختلاف مستويات طول الاختبار (15، 30، 60) فقرة،

وفقا للنموذج أحادي المعلم، حيث كانت قيم مستوى الاحتمالية غير دالـــة إحـــصائيا (p = 0.258, 0.676, 0.23) عند مستوى الدلالة (α = 0.05)، وكان التطابق أكبر ما يمكن عندما كان طول الاختبار مساوٍ 30 فقرة (Bollen & Long, 1993).

وأظهرت النتائج تطابقا في حالة توزيع القدرة حينما كان ملتويا التواءً موجبا عندما كان عدد الفقرات 15 فقرة (p = 0.378) وعند مستوى الدلالة (α= 0.05)، في حين أظهرت النتائج الأخرى وجود فروق ذات دلالة إحصائية لقيم (χ^2) في الحـــالات الأخـــرى مـــن مستوى طول الاختبار (30، 60) لنفس شكل التوزيع. والملتوِ التواءً سالبا علـــى اخـــتلاف طول الاختبار، وفقا للنموذج أحادي المعلم وهذا يشير لعدم تطابق بين النموذج والبيانـــات المولدة.

وعندما تم توليد البيانات وفقا للنموذج الثنائي المعلمة كانت النتائج تشير إلى عـــدم تطابق البيانات للنموذج لجميع حالات مستوى طول الاختبار(15، 30، 60)، وشكل توزيع القدرة (طبيعي، ملتو التواء موجب، ملتو التواء سال)، حيث تـــشير مـــستوى الاحتماليـــة لوجود فروق ذات دلالة إحصائية عند كل قيمة من قيم المتغير المستقل.

أما عند توليد البيانات وفقا للنموذج الثلاثي المعلمة كانت النتائج تشير إلى تطـــابق البيانات للنموذج عند شكل توزيع القدرة الملتوِ التواءً موجبا حينما كان عـــدد الفقـــرات 30 فقرة من مستوى طول الاختبار، حيث بلغت مستوى الاحتمالية (p = 0.095) وهذا يشير لعدم وجود فروق ذات دلالة إحصائية وهذا يؤكد على جودة المطابقة، فـــي حـــين تـــشير الحالات الأخرى إلى عدم تطابق البيانات للنموذج عند مستويات طول الاختبار(15، 60)، وشكل توزيع القدرة (طبيعي، ملتو التواء سالب)، حيث تشير مستوى الاحتماليـــة لوجـــود فروق ذات دلالة إحصائية عند كل قيمة من قيم المتغير المستقل وعند مستوى الدلالة (α =0.05).

النتائج المتعلقة بالمؤشرات RMSEA، RMR:

وبنفس الطريقة السابقة تم توليد بيانات بواقع (1000) مفحوص لكافة متغيرات الدراسة(طول الاختبار (15، 30، 60) فقرة، شكل توزيع القدرة (ملتو التواء موجب وملتو التواء سالب، طبيعي)، وفق نماذج استجابة الفقرة (أحادي المعلم، ثنائي المعلم، ثلاثي المعلم))، تم تحليل البيانات المولدة باستخدام برمجية AMOS 21 واستخراج مؤشرات جودة المطابقة (Fit Indicators)، باستخدام التحليل العاملي التوكيدي (Confirmatory Factor Analyzes) للكشف عن أحادية البعد، وتم التوصل لكافة النتائج المتعلقة بتلك المؤشرات وهي:

1- الجذر التربيعي لمتوسط مربع الخطأ التقريبي (RMSEA): يؤكد بلنتر وبونت (Bentler & Bonett, 1980) أن هذا المؤشر من أهم مؤشرات جودة المطابقة، وإن كانت قيمته محصورة بين (0.05 ، 0.08) دل على أن النموذج يطابق بيانات العينة، وإن كانت قيمته 0.03 فأقل دل على أن النموذج يطابق البيانات بصورة كبيرة، وإن كان صفرا دل على التطابق التام، وقد تم اعتماد علامة قطع متفق عليها وفقا للأدب النظري بحيث تكون قيمة المؤشر أقل من 0.01 (RMSEA value < 0.01) (MacCallum, Browne and Sugawara 1996).

2- مؤشر الجذر التربيعي لمربعات الأخطاء Root Mean Square Residual (RMR): وهو من المؤشرات الهامة جدا وكلما كانت قيمته أقل وقريبة من الصفر تعطي أكثر تطابقا، أما إذا زادت قيمته عن 0.08 يتم رفض النموذج (أي بمعنى عدم تطابق). والقيمة التي تم اعتمادها للمؤشر (RMR) وفقا للأدب النظري (RMR < 0.05) Bentler & Bonett, 1980 (Hu & Bentler, 1999). ويبين الجدول 9 عرضا تفصيليا للنتائج:

جدول 9: قيم المؤشر (RMSEA, RMR) وفقا لنماذج استجابة الفقرة باختلاف طول الاختبار وشكل توزيع القدرة

النموذج	توزيع القدرة	عدد الفقرات	RMSEA	RMR
أحادي المعلم	ملتو التواء سالب	15	0.007	0.004
		30	0.006	0.005
		60	0.007	0.005
	طبيعي	15	0.003	0.005
		30	0	0.006
		60	0.002	0.007
	ملتو التواء موجب	15	0.002	0.004
		30	0.006	0.005
		60	0.008	0.005
ثنائي المعلم	ملتو التواء سالب	15	0.014	0.005
		30	0.012	0.005
		60	0.012	0.006
	طبيعي	15	0.013	0.007
		30	0.009	0.007
		60	0.005	0.007
	ملتو التواء موجب	15	0.026	0.008
		30	0.014	0.005
		60	0.014	0.006
ثلاثي المعلم	ملتو التواء سالب	15	0.019	0.004
		30	0.015	0.004
		60	0.015	0.006
	طبيعي	15	0.008	0.005
		30	0.006	0.006
		60	0.006	0.007
	ملتو التواء موجب	15	0.005	0.007
		30	0.003	0.007
		60	0.003	0.008

يتضح من الجدول 9 أن قيم المؤشر RMSEA حققت تطابقا بين البيانات المولـــدة والنموذج المتوقع وفقا لنماذج استجابة الفقرة أحادية البعد، وعند طول الاختبـــار (15، 30، 60)، وشكل توزيع القدرة (ملتو التواء سالب وطبيعي وملتو التواء موجب). وقد حققت تطابقا

تاما عندما توزعت القدرة بشكل طبيعي عند طول الاختبار(30) فقرة، حيث كانت قيمة المؤشر (RMSEA=0) وهذا يشير إلى التطابق التام (A Close-Fitting Model).

في حين أن النتائج الأخرى تشير إلى تطابق البيانات للنموذج عندما توزعت القدرة بشكل طبيعي عند طول الاختبار (30، 60) فقرة. وفقا لنموذج ثنائي المعلمة، أما وفقا لنموذج ثلاثي المعلمة كان التطابق عندما توزعت القدرة بشكل طبيعي وبشكل ملتو التواء موجب وعند كافة مستويات طول الاختبار. أي أنه حقق نتائج ملموسة وفقا للمعيار المتفق عليه (RMSEA value < 0.01). وبالرجوع إلى الجدول 9 ودراسة قيم المؤشر RMR يتبين انه حقق تطابقا كبيرا تبعا لعلامة القطع المتفق عليها؛ فجميع قيم المؤشر RMR حققت المعيار (RMR < 0.05) بين النموذج والبيانات بصورة كبيرة وواضحة عند كافة المتغيرات : طول الاختبار (15، 30، 60) فقرة، وشكل توزيع القدرة (ملتو التواءً سالبا، طبيعي، ملتو التواءً موجبا) وفقا لنماذج استجابة الفقرة (أحادي وثنائي وثلاثي المعلمة).

ولكن يجب أن لا يغيب عن الأنظار أن هذا المؤشر يتأثر بحجم العينة بصورة كبيرة وقد يحقق تطابقا عند زيادة حجم العينة، وهذا من شأنه أن يقلل من جودة المطابقة، إذ أن الحجوم الكبيرة تسبب نقصان في قيمة المؤشر، ولتحييد هذا الأثر تم ضبط حجم العينة (1000) مفحوص لكافة مستويات المتغير المستقل، بالإضافة لذلك لا يمكن الحكم على جودة المطابقة من خلال ذلك المؤشر بصورة انفرادية (Kenny and McCoach,2003).

النتائج المتعلقة بمؤشرات المطابقة المتزايدة Incremental Fit Indexes (Bollen, 1986):

وبنفس الطريقة السابقة تم توليد بيانات بواقع (1000) مفحوص لكافة متغيـرات الدراسة؛ طول الاختبار (15، 30، 60) فقرة، شكل توزيع القدرة (ملتو التواءً موجبا وملتو التواءً سالبا، طبيعي)، وفق نماذج استجابة الفقرة (أحادي المعلم، ثنائي المعلم، ثلاثي المعلم)، تم تحليل البيانات المولدة باستخدام برمجية AMOS 21 واسـتخراج مؤشـرات جـودة المطابقة (Fit Indicators)، باستخدام التحليل العاملي التوكيدي للكشف عن أحادية البعـد، وتم التوصل لكافة النتائج وهي:

1- مؤشر المطابقة المعياري Normed Fit Index (NFI).

يعد من المؤشرات المتشددة لدرجات الحرية، وعلامة القطع وفقا للأدب النظري هي (value > 0.8) (Kenny, Kaniskan, and McCoach ,2011).

2- مؤشر المطابقة المقارن Comparative Fit Index (CFI).

من المؤشرات المعيارية الجيدة، وعلامة القطع المعتمدة وفقا للأدب النظـري هـي (value > 0.9) لقبول النموذج ومطابقته للبيانات، وهو لا يعتمد على حجم العينة لـذا يفضل على العديد من المؤشرات الأخرى، وقد أوصى به العديـد مـن البـاحثين (Gerbing and Anderson ,1993).

3- مؤشر توكر لويس Tucker-Lewis Index (TLI).

وهو من المؤشرات المعيارية الجيدة غير المتشددة، وعلامة القطع المعتمدة من قبـل الباحثة وفقا للأدب النظري (value > 0.8) لقبول النموذج، وهو لا يعتمد أيضا على حجم العينة ويفضل على العديد من المؤشرات الأخرى، إلا انه عند الحجوم الصغيرة

تقل قيمته، وهو يرتبط بصورة مباشرة بمؤشر CFI ويسجل نتائج أقل، وقد أوصى به العديد من الباحثين. Gerbing and Anderson ,1993, Kenny, Kaniskan, and (McCoach ,2011).

4- مؤشر المطابقة المتزايد Incremental fit index (IFI).

وهو من المؤشرات المعيارية المتشددة، وعلامة القطع المعتمدة هي (value > 0.9) وتشير القيم المرتفعة منه (أكبر من 0.90) إلى تطابق البيانات مع النموذج.

5- مؤشر المطابقة النسبي (Relative fit index (RFI.

من المؤشرات المعيارية والذي يتم الحصول عليه من مؤشر NFI بإجراء تعديل على درجات الحرية لذا يعتبر من المؤشرات المتشددة جدا. وتشير القيم المرتفعة منه (أكبر من 0.90) إلى تطابق البيانات مع النموذج أما ان كانت أكبر من 0.95 تشير الى أفضل تطابق. وتشير القيمة واحد صحيح الى التطابق التام(Gerbing and Anderson, 1993) . وتم الحصول على النتائج وجدولتها جميعها بجدول واحد يضم كافة المؤشرات السابقة، كونها من عائلة واحدة تندرج أسفل مؤشرات المطابقة المتزايدة، والجدول 10 يوضح النتائج بصورة تفصيلية:

جدول 10: قيمة المؤشرات (NFI , RFI, IFI, TLI, CFI) للبيانات المولدة وفقــا لنمــاذج استجابة الفقرة باختلاف طول الاختبار وشكل توزيع القدرة

النموذج	توزيع القدرة	عدد الفقرات	NFI	RFI	IFI	TLI	CFI
أحادي المعلم	ملتو التواء سالب	15	0.698	0.648	0.921	0.903	0.917
		30	0.443	0.401	0.850	0.827	0.839
		60	0.170	0.141	0.522	0.467	0.485
	طبيعي	15	0.738	0.694	0.986	0.982	0.985
		30	0.478	0.439	1.000	1.000	1.000
		60	0.198	0.170	0.957	0.949	0.950
	ملتو التواء موجب	15	0.762	0.723	0.994	0.993	0.994
		30	0.446	0.405	0.844	0.820	0.832
		60	0.165	0.136	0.459	0.403	0.423
ثنائي المعلم	ملتو التواء سالب	15	0.629	0.567	0.771	0.722	0.762
		30	0.391	0.346	0.618	0.571	0.601
		60	0.137	0.106	0.271	0.218	0.245
	طبيعي	15	0.684	0.631	0.820	0.783	0.814
		30	0.428	0.386	0.722	0.685	0.707
		60	0.194	0.165	0.697	0.655	0.666
	ملتو التواء موجب	15	0.515	0.434	0.579	0.499	0.570
		30	0.369	0.322	0.548	0.497	0.531
		60	0.124	0.094	0.224	0.173	0.201
ثلاثي المعلم	ملتو التواء سالب	15	0.554	0.479	0.661	0.591	0.650
		30	0.313	0.263	0.462	0.401	0.442
		60	0.110	0.079	0.187	0.137	0.167
	طبيعي	15	0.671	0.616	0.888	0.862	0.882
		30	0.445	0.404	0.854	0.832	0.843
		60	0.177	0.149	0.617	0.566	0.581
	ملتو التواء موجب	15	0.664	0.609	0.941	0.926	0.937
		30	0.399	0.355	0.943	0.932	0.937
		60	0.185	0.156	0.855	0.828	0.834

يتضح من الجدول 10 أن مؤشر NFI (Normed Fit Index) قد تراوحت قيمته بين (0.11) وبين (0.762)، وقد حقق نتائج مقبولة فقط عندما توزعت البيانات بشكل ملتو التواءً موجباً عند طول الاختبار (15 فقرة) وفقا لنموذج أحادي البعد، حيث كانت قيمـــة المؤشـــر ضمن الحد المقبول (0.762) مقارنة مع علامة القطع المتفق عليها (value > 0.8).

أما بقية النتائج المتعلقة بالمؤشر NFI لم تحقق أي تطابق يذكر بين البيانات والنموذج المتوقع عند كافة مستويات المتغير المستقل من طول الاختبار (15، 30، 60) فقرة، وشكل توزيع القدرة (ملتو التواء سالب، طبيعي، ملتو التواء موجب)، وفقا لنماذج استجابة الفقرة المختلفة (أحادي المعلم، ثنائي المعلم، ثلاثي المعلم).

ولدى دراسة المؤشر Relative fit index RFI (وهو من المؤشرات المتشددة جدا)؛ لم يحقق أي نتائج تذكر من جودة المطابقة ضمن حدود القطع المتفق عليها (value > 0.9)، وعند جميع مستويات المتغير المستقل من طول الاختبار (15، 30، 60) فقرة وشكل توزيع القدرة (ملتو التواء سالب، طبيعي، ملتو التواء موجب)، وفقا لنماذج استجابة الفقرة المختلفة (أحادي المعلم، ثنائي المعلم، ثلاثي المعلم)، والنتائج السابقة تؤكد أن هذا المؤشر من المؤشرات المتشددة، حيث لم يتم الحصول على أي تطابقٍ يذكر.

وبدراسة نتائج المؤشر IFI (Incremental fit index) يلاحظ أنه حقق نتائج ملموسة من جودة المطابقة ضمن حدود القطع المتفق عليها (value >0.9)، وكان ذلك عندما توزعت القدرة بشكل طبيعي وعند كافة مستويات المتغير المستقل (طول الاختبار)، وحينما توزعت القدرة بشكل ملتو التواء سالب عند مستوى (15، 30) فقرة من طول الاختبار، وحينما توزعت بشكل ملتو التواء موجب عند مستوى (15) فقرة من طول الاختبار وفقا للنموذج أحادي المعلم. كما وأظهرت النتائج تطابقا تاما (IFI=1)، عند مستوى 30 فقرة من طول الاختبار وتوزيع طبيعي لمستوى شكل القدرة وفقا لنموذج أحادي المعلم. ولم يقدم المؤشر IFI أي نتائج تذكر من جودة المطابقة ضمن حدود القطع المتفق عليها وفقا لنموذج ثنائي المعلم.

أما البيانات المولدة وفقا للنموذج ثلاثي المعلم لنماذج استجابة الفقرة قدم المؤشر IFI فيها نتائج ملموسة عندما توزعت القدرة بشكل طبيعي وعند طول الاختبار (15، 30) فقرة، وعندما توزعت القدرة بشكل متلو التواء موجب لكافة أطوال الاختبار (15، 30، 60) فقرة.

وبدراسة نتائج المؤشر (Tucker-Lewis Index TLI) يتبين أنـــه حقـــق نتـــائج ملموسة ضمن حدود القطع المتفق عليها (value > 0.8)، وذلك عندما توزعت القدرة بشكل طبيعي وعند كافة مستويات المتغير المستقل (طول الاختبار)، وحينما توزعت القدرة بشكل ملتو التواء سالب عند مستوى (15، 30) فقرة، وحينما توزعت بشكل ملتو التواء موجـــب عند مستوى (15، 30) فقرة وفقا للنموذج أحادي المعلم. كما وأظهرت النتائج حالة التطابق التام (TLI=1) عند مستوى 30 فقرة من طول الاختبار وشكل التوزيع طبيعي لمستوى شكل القدرة وفقا لنموذج أحادي المعلم.

وقدم المؤشر TLI نتائج واضحة وفقا لنموذج ثنائي المعلم لنماذج استجابة الفقـــرة عند مستوى طبيعي من شكل توزيع القدرة وعند مستوى (15) فقرة من طول الاختبـــار، أما وفقا للنموذج ثلاثي المعلم قدم المؤشر السابق نتائج ملموسة في المستوى الطبيعي لشكل القدرة (15، 30) فقرة لمستوى طول الاختبار، وشكل التوزيع ألمتلو التواء موجب عنـــد جميع مستويات طول الاختبار (15، 30، 60).

وبدراسة نتائج المؤشر Comparative Fit Index) CFI) يلاحظ أنـــه حقـــق نتـــائج ملموسة ضمن حدود القطع المتفق عليها (value > 0.9)، عندما توزعت القدرة بشكل طبيعي وعند كافة مستويات المتغير المستقل (طول الاختبار)، وحينما توزعت القدرة بشكل ملتو التواء سالب عند مستوى (15) فقرة، وحينما توزعت بشكل ملتو التواء موجب عنـــد مـــستوى (15) فقرة وفقا للنموذج أحادي المعلم. كما وأظهرت النتائج تطابقا تاما (CFI=1)، عند مستوى 30 فقرة

من طول الاختبار وتوزيع طبيعي لمستوى شكل القدرة وفقا لنموذج أحادي المعلم، ولم يقدم المؤشر CFI أي نتائج تذكر وفقا لنموذج ثنائي المعلم، أما وفقا للنموذج ثلاثي المعلـــم قـــدم المؤشر نتائج ملموسة في المستوى الطبيعي لشكل القدرة (15) فقرة، والمـــستوى ألمتلـــو التواء موجب عند مستويات طول الاختبار (15، 30) فقرة. يتبين من خلال الجدول 10 أن قيم المؤشر IFI كانت أكبر من قيم المؤشر CFI، ويليه المؤشر TLI، إلا أن قيمـــة القطـــع للمؤشر الأخير كانت أقل منهما الاثنين مما حقق تطابقا أكثر لدى المؤشر TLI.

قيم المؤشرات المعدلة لدرجات الحرية (PNFI ،PCFI):

1. (PNFI :Parsimony Adjustment to the NFI). مؤشر المطابقة المعياري المعدل
2. (PCF:Parsimony Adjustment to the CFI) . مؤشر المطابقة المقارن المعدل

وهي معدلة لمؤشرات المطابقة المعيارية بـــدرجات الحريـــة للنمـــوذج الـــصفري والنموذج المتوقع وهي أكثر تشددا لدرجات الحرية. وتعتبر هـــذه المؤشـــرات (PNFI,) PCFI اشتقاقا للمؤشرات السابقة، وأن المؤشر PCFI يتأثر بمستوى طول الاختبار، وان علامة القطع المتفق عليها أكبر من 0.9 لكلا المؤشرين، ويوضح الجدول الآتـــي النتـــائج بصورة تفصيلية (Gerbing and Anderson ,1993):

جدول 11: قيمة المؤشرات (PNFI, PCFI) وفقا لنماذج استجابة الفقرة باختلاف طول الاختبار وشكل توزيع القدرة

النموذج	توزيع القدرة	عدد الفقرات	PNFI	PCFI
أحادي المعلم	ملتو التواء سالب	15	0.599	0.786
		30	0.412	0.781
		60	0.165	0.468
	طبيعي	15	0.633	0.844
		30	0.445	0.931
		60	0.192	0.918
	ملتو التواء موجب	15	0.653	0.852
		30	0.416	0.775
		60	0.159	0.408
ثنائي المعلم	ملتو التواء سالب	15	0.539	0.653
		30	0.364	0.559
		60	0.132	0.236
	طبيعي	15	0.586	0.698
		30	0.399	0.658
		60	0.187	0.644
	ملتو التواء موجب	15	0.442	0.489
		30	0.343	0.495
		60	0.120	0.194
ثلاثي المعلم	ملتو التواء سالب	15	0.475	0.557
		30	0.292	0.412
		60	0.106	0.161
	طبيعي	15	0.575	0.756
		30	0.414	0.785
		60	0.171	0.561
	ملتو التواء موجب	15	0.570	0.803
		30	0.372	0.872
		60	0.179	0.806

يتضح من الجدول 11 أن قيم المؤشر PNFI لم تحقق أي نتائج مطابقة عند جميع مستويات المتغير المستقل من طول الاختبار (15، 30، 60) فقرة وشكل توزيع القدرة

(ملتو التواءً سالبا، طبيعي، ملتو التواءً موجبا)، وفقا لنماذج استجابة الفقرة (أحادي المعلم، ثنائي المعلم، ثلاثي المعلم)، بالرغم من تحسن بعض القيم وفقا للنمــوذج أحــادي المعلــم، وتظهر النتائج تأثر هذا المؤشر بمستوى التغير في طول الاختبار بشكل واضح.

أما المؤشر PCFI قد أظهر نتائج ملموسة عندما توزعت القدرة بشكل طبيعي عنــد مستوى (30، 60) فقرة من طول الاختبار، وعندما كان توزيع البيانات ملتو التواء موجب عند مستوى (15) فقرة من طول الاختبار وفقا للنموذج أحادي المعلم. ولم يقدم المؤشــر PCFI أي نتائج تذكر وفقا لنموذج ثنائي المعلم، ووفقا للنموذج ثلاثي المعلم قــدم المؤشــر السابق نتائج ملموسة في عندما توزعت القدرة بشكل ملتو التواء موجب عند مستوى (30) فقرة من طول الاختبار.

مؤشرات المطابقة المطلقة (Absolute Fit Indexes) :

ومن هذه المؤشرات :(Bentler & Bonett, 1980) (Hewitt, et al., 2004)

1- مؤشر حسن المطابقة Goodness of Fit Index (GFI).

2- مؤشر حسن المطابقة المصحح بدرجات الحرية (AGFI).

3- مؤشر حسن المطابقة المتشدد parsimony goodness of fit index (PGFI).

وتشير القيم المرتفعة لهذه المؤشرات (AGFI) (GFI) إلى تطابق أفضل مع النموذج وعلامة القطع المتفق عليها (أكبر من 0.90)، ويكون التطابق التام في حالة الواحد الصحيح لقيمة المؤشر. ومن المعلوم أن المؤشرات السابقة تتأثر كثيرا بحجوم العينات، إذ أن العينات الكبيرة تقدم نتائج مطابقة أفضل مما يشكك بمصداقيتها، وكثيرا من الباحثين لا ينــصحون باستخدامها لوحدها، لضمان صدق جودة المطابقة، كما أن المؤشر AGFI قد لا يعطــي أي

نتائج عند استخدام عينات بحجوم صغيرة، ويتأثر بمستوى التغير في طول الاختبار، إذ تزداد قيمة ذاك المؤشر بزيادة طول الاختبار. ويوضح الجدول 12 النتائج بصورة تفصيلية: (Sharma, Mukherjee, Kumar, & Dillon, 2005).

جدول 12: قيمة المؤشرات (PGFI, AGFI, GFI) للبيانات المولدة وفقا لنماذج استجابة الفقرة باختلاف طول الاختبار وشكل توزيع القدرة

النموذج	توزيع القدرة	عدد الفقرات	GFI	AGFI	PGFI
أحادي المعلم	ملتو التواء سالب	15	0.985	0.980	0.739
		30	0.969	0.964	0.844
		60	0.930	0.925	0.869
	طبيعي	15	0.987	0.983	0.741
		30	0.973	0.969	0.848
		60	0.942	0.938	0.881
	ملتو التواء موجب	15	0.988	0.984	0.741
		30	0.968	0.964	0.843
		60	0.926	0.920	0.865
ثنائي المعلم	ملتو التواء سالب	15	0.976	0.968	0.732
		30	0.955	0.948	0.832
		60	0.901	0.894	0.842
	طبيعي	15	0.977	0.970	0.733
		30	0.962	0.956	0.838
		60	0.936	0.932	0.875
	ملتو التواء موجب	15	0.947	0.930	0.710
		30	0.948	0.940	0.826
		60	0.888	0.880	0.829
ثلاثي المعلم	ملتو التواء سالب	15	0.967	0.956	0.725
		30	0.942	0.934	0.821
		60	0.877	0.869	0.820
	طبيعي	15	0.984	0.978	0.738
		30	0.969	0.964	0.844
		60	0.934	0.930	0.873
	ملتو التواء موجب	15	0.986	0.982	0.740
		30	0.972	0.968	0.846
		60	0.941	0.936	0.879

يتضح من خلال الجدول 12 أن قيم المؤشر (GFI) والمؤشر (AGFI) تشير إلى تطابق كبير للبيانات مع النموذج، عند جميع مستويات المتغير المستقل من طول الاختبار (15، 30، 60) فقرة وشكل توزيع القدرة (ملتو التواء سالبا، طبيعي، ملتو التواء موجبا)، وفقا

لنماذج استجابة الفقرة (أحادي المعلم، ثنائي المعلم، ثلاثي المعلم)، وهذا التطابق يؤكد أن حجم العينة الكبير كان له أثرا واضحا على النتائج. حيث أكد كيني وماكوش (Kenny & McCoach, 2003) أن زيادة حجم العينة تؤثر في جودة المطابقة، نماذج بحجوم عينات كبيرة تؤدي لقيم أقل لمؤشرات جودة المطابقة.

أما المؤشر PGFI قدم نتائج ضعيفة، فهو من المؤشرات المتشددة والتي تتأثر بطول الاختبار، وقد أعطى تطابقا في الحالات الآتية:

- عندما كان عدد الفقرات 60 فقرة لجميع مستويات شكل القدرة (ملتو سالب، طبيعي، ملتو موجب) وفقا للنموذج أحادي المعلم.
- عندما كان عدد الفقرات 60 فقرة وتوزعت البيانات بشكل طبيعي وفقا للنموذج ثنائي المعلم.
- عندما كان عدد الفقرات 60 فقرة وتوزعت بشكل طبيعي وبشكل ملتو التواء موجب وفقا للنموذج ثلاثي المعلم.

مؤشر هولتر لجودة المطابقة بمستوى (0.05)، وبمستوى (0.01) Hoelter's (1983)`critical N' for a significance level of .05:

يقارن هولتر (Hoelter's ,1983) قيمة المؤشر في التحليل بالحجم المسموح به وهو يناظر حجم العينة في التحليل الإحصائي المستخدم، فإن كانت قيمة مؤشر هولتر أكبر من الحجم المسموح به فهذا يعني تطابق البيانات للنموذج، وان كان أقل يتم رفض النموذج وتكون المقارنة عادة على مستوى الدلالة (α= 0.05, 0.01). وقد أشار هيو وبلنتر (Hu and Bentler

(1998, لحساسية هذا المؤشر لحجوم العينات الكبيرة (N > 200)، بحيث يعطي قيمـا ذات دلالة إحصائية، وبالتالي لا يوصى باستخدامه.

وقد تم المقارنة بحجم العينة مضروبا بعـدد مـرات التكـرار Number of Replication (5*1000=5000)، وبالتالي أي قيمة أقل من العدد 5000 فهي غيـر مطابقـة للنموذج. وفيما يلي النتائج بصورة تفصيلية:

جدول 13: المؤشرات (HOELTER, 0.05) وفقا لنماذج استجابة الفقرة باختلاف طول الاختبار وشكل توزيع القدرة

النموذج	توزيع القدرة	عدد الفقرات	HOELTER (0.05)	HOELTER (0.01)
أحادي المعلم	ملتو التواء سالب	15	4457	4655
أحادي المعلم	ملتو التواء سالب	30	2761	2884
أحادي المعلم	ملتو التواء سالب	60	2022	2112
أحادي المعلم	طبيعي	15	5333	5570
أحادي المعلم	طبيعي	30	2928	3058
أحادي المعلم	طبيعي	60	4136	4320
أحادي المعلم	ملتو التواء موجب	15	5454	5696
أحادي المعلم	ملتو التواء موجب	30	1269	1325
أحادي المعلم	ملتو التواء موجب	60	4872	5089
ثنائي المعلم	ملتو التواء سالب	15	4524	4619
ثنائي المعلم	ملتو التواء سالب	30	3168	3235
ثنائي المعلم	ملتو التواء سالب	60	2463	2515
ثنائي المعلم	طبيعي	15	5338	5451
ثنائي المعلم	طبيعي	30	3743	3823
ثنائي المعلم	طبيعي	60	4539	4636
ثنائي المعلم	ملتو التواء موجب	15	4474	4569
ثنائي المعلم	ملتو التواء موجب	30	2728	2786
ثنائي المعلم	ملتو التواء موجب	60	5051	5158
ثلاثي المعلم	ملتو التواء سالب	15	4162	4205
ثلاثي المعلم	ملتو التواء سالب	30	2945	2976
ثلاثي المعلم	ملتو التواء سالب	60	2383	2407
ثلاثي المعلم	طبيعي	15	5069	5121
ثلاثي المعلم	طبيعي	30	4590	4638
ثلاثي المعلم	طبيعي	60	4441	4487
ثلاثي المعلم	ملتو التواء موجب	15	3934	3974
ثلاثي المعلم	ملتو التواء موجب	30	2603	2630
ثلاثي المعلم	ملتو التواء موجب	60	4929	4980

يتضح من خلال الجدول 13 أن مؤشر (0.05) HOELTER حقق نتائج مقبولة عندما توزعت البيانات توزيعا طبيعي عند طول اختبار (15 فقرة) وفقا لنموذج أحـــادي المعلـــم، حيث كانت قيمة المؤشر ضمن الحد (5333)، وحققت نتائج مقبولة عندما توزعت البيانـــات توزيعا ملتويا موجبا عند مستوى طول اختبار (15 فقرة)، وفقا لنموذج أحادي المعلم.

وفي النموذج ثنائي المعلم لم تحقق أي تطابق عندما توزعت القدرة بـــشكل ملتـــو التواء سالب، في حين أنها حققت تطابقا مقبولا عندما توزعت البيانات بشكل طبيعي عنـــد مستوى 15 فقرة من طول الاختبار، وعندما توزعت بشكل ملتو التواء موجب عند مستوى 60 فقرة من طول الاختبار.وفي النموذج ثلاثي المعلم لم يحقق المؤشر أي نتائج تذكر إلا في حالة واحدة وهي عندما توزعت البيانات بشكل طبيعي وعند مستوى 15 فقرة من طول الاختبار.

ولدى دراسة المؤشر (0.01) HOELTER يلاحظ أنـــه أبـــدى توافقـــا بينـــه وبـــين (0.05) HOELTER، الا أن قيمه كانت أكبر عند جميع مستويات المتغير المـــستقل، وهـــذا يزيد من فرصة جودة المطابقة، حيث أنه أبدى تحسينا ملموسا وأبدى تطابقا مقبولا عنـــدما توزعت البيانات بشكل ملتو التواء موجب عند مستوى 60 فقرة من طول الاختبـــار وفقـــا للنموذج أحادي المعلم، في حين أن المؤشر (0.05) HOELTER لم يعطي توافقا عند تلك الحالة.

ثالثاً النتائج المتعلقة بالإجابة عن أسئلة الدراسة:

أولا: النتائج المتعلقة بأثر طول الاختبار على فاعلية المؤشرات المستخدمة في الكـــشف عن افتراض أحادية البعد.

للإجابة عن السؤال الأول والذي ينص على: هل تختلف فاعلية مؤشرات الكـــشف عن تحقق افتراض أحادية البعد في النموذج البنائي المستخدم باختلاف طول الاختبار (15، 30، 60) فقرة؟

تم حساب أثر طول الاختبار على فاعلية المؤشر المستخدم في الكشف عن افتــرا أحادية البعد، من خلال حساب χ^2 لحسن المطابقة، وتحييد أثر كل من نوع النموذج وشــكل توزيع القدرة. و الجدول 17 يوضح النتائج:

جدول 14: قيمة الإحصائي (χ^2) بتغيير طول الاختبار وتحييد أثر شكل توزيع القدرة ونوع النموذج المستخدم

significance	Df	χ^2	Standardized Residual الباقي المعياري	Residual الباقي	Expected Frequencies التكرارات	Observed Frequencies	INDICOTER المؤشر	الرتبة	عدد الفقرات
0.087	10	16.481	1.85	4.09	4.91	9	AGFI	1	15
			1.85	4.09	4.91	9	GFI	2	
			1.85	4.09	4.91	9	RMR	3	
			0.04	0.09	4.91	5	RMSEA	4	
			0.04	0.09	4.91	5	TLI	5	
			-0.41	-0.91	4.91	4	CFI	6	
			-0.41	-0.91	4.91	4	IFI	7	
			-0.86	-1.91	4.91	3	HOELTER 0.01	8	
			-1.31	-2.91	4.91	2	HOELTER 0.05	9	
			-1.31	-2.91	4.91	2	PGFI	10	
			-1.31	-2.91	4.91	2	χ^2 /DF	11	
			0.00	0.00	54.00	54	الكلي		
0.006	11	26.412	2.30	4.75	4.25	9	AGFI	1	30
			2.30	4.75	4.25	9	GFI	2	
			2.30	4.75	4.25	9	RMR	3	
			0.85	1.75	4.25	6	RMSEA	4	
			0.36	0.75	4.25	5	TLI	5	
			-1.09	-2.25	4.25	2	CFI	6	
			-1.09	-2.25	4.25	2	HOELTER 0.01	7	
			-1.09	-2.25	4.25	2	HOELTER 0.05	8	
			-1.09	-2.25	4.25	2	IFI	9	
			-1.09	-2.25	4.25	2	PGFI	10	
			-1.09	-2.25	4.25	2	χ^2 /DF	11	
			-1.58	-3.25	4.25	1	PCFI	12	
			0.00	0.00	51.00	51	الكلي		
0.001	11	30.211	3.28	5.83	3.17	9	RMR	1	60

significance	Df	χ^2	Standardized Residual الباقي المعياري	Residual الباقي	Expected Frequencies التكرارات	Observed Frequencies	INDICOTER المؤشر	الرتبة	عدد الفقرات
			2.15	3.83	3.17	7	GFI	2	
			1.59	2.83	3.17	6	AGFI	3	
			1.59	2.83	3.17	6	RMSEA	4	
			-0.66	-1.17	3.17	2	TLI	5	
			-0.66	-1.17	3.17	2	χ^2 /DF	6	
			-1.22	-2.17	3.17	1	CFI	7	
			-1.22	-2.17	3.17	1	HOELTER 0.01	8	
			-1.22	-2.17	3.17	1	HOELTER 0.05	9	
			-1.22	-2.17	3.17	1	IFI	10	
			-1.22	-2.17	3.17	1	PGFI	11	
			-1.22	-2.17	3.17	1	PCFI	12	
			0.00	0.00	38.00	38	الكلي		

بالرجوع إلى الجدول 14 يتبين أن قيمة مؤشر χ^2 لحسن المطابقــة ($p = 0.006$) كانت دالة إحصائية عند مستوى الدلالة ($\alpha = 0.05$) بدرجات حرية 11 عند طول اختبــار 30 فقرة، لصالح المؤشرات AGFI، GFI، RMR. كما حققت دلالة إحصائية ($p = 0.001$) وبدرجات حرية 11عند مستوى 60 فقرة، ولصالح RMR، GFI. ولم يقدم المؤشــر χ^2 أي تطابق عند طول 15 فقرة من الاختبار، حيث أن قيمة المؤشر χ^2 لحسن المطابقــة كانــت غير دالة إحصائيا عند مستوى الدلالة ($\alpha = 0.05$) للمؤشرات جميعها.

وتم اعتماد قيمة الباقي المعياري للحكم على جودة المطابقة، حيث أشــار هنكلــز، وريزما و جورس (Hinkle, Wiersma, & Jurs, 1988) إلى اعتماد قيمة الباقي المعيــاري كمعيار للحكم، بحيث تكون القيمة (-2) أو أقل ، (+2) أو أكبر (Standardized Residual would be -2 and +2)، لاعتمادها كقيمة مقبولة لجودة المطابقة.

وهذا يتضح من الجدول 14، حيث تعد القيم الموجبة في حساب الباقي المعياري لقيم المؤشرات مؤشراً على فاعلية ذاك المؤشر. كما ينصح في استخدامها في الكشف عن افتراض أحادية البعد. ويظهر من خلال الجدول أن المؤشرات التالية (GFI, RMR, AGFI) تصلح للكشف عن أحادية البعد في حالة تغيير طول الاختبار. هذا يعني أن تلك المؤشرات لا تتأثر كثيرا بتغيير طول الاختبار، وهي المناسبة في الكشف عن افتراض أحادية البعد لدى تغيير طول الاختبار. كما أظهرت النتائج أن طول الاختبار المفضل في الكشف عن أحادية البعد هو 30 فقرة. في حين أن القيم السالبة للباقي المعياري منها لا ينصح في استخدامها (PCFI ،PGFI ،IFI ،HOELTER 0.05 ،HOELTER 0.01 ،CFI). مما يعني أن تلك المؤشرات تتأثر بطول الاختبار.

ويمكن القول أن القيم الموجبة في حساب الباقي المعياري للمؤشرات أثبتت فاعليتها وحققت تطابقاً بين البيانات والنموذج المقترح وأبدت فاعليتها في الكشف عن افتراض أحادية البعد، ويمكن أن يوصى في استخدامها للكشف عن افتراض أحادية البعد لدى تغيير طول الاختبار.

ثانيا: النتائج المتعلقة بأثر شكل توزيع القدرة على فاعلية المؤشرات المستخدمة في الكشف عن افتراض أحادية البعد.

للإجابة عن السؤال الثاني والذي ينص على: هل تختلف فاعلية مؤشرات الكشف عن تحقق افتراض أحادية البعد في النموذج البنائي المستخدم باختلاف شكل توزيع القدرة (التوزيع الطبيعي، ملتوٍ التواءً موجباً، ملتوٍ التواءً سالباً)؟

لتحقيق ذاك الغرض تم حساب الاختبار الإحصائي χ^2 لحسن المطابقة لكافة المؤشرات لدراسة أثر تغيير شكل توزيع القدرة، بغض النظر عن نوع المتغير المستقل المستخدم (طول الاختبار، ونوع النموذج)، وكانت النتائج موضحة في الجدول 16:

جدول 15: قيمة الإحصائي (χ^2) بتغيير شكل توزيع القدرة وتحييد أثر طول الاختبار ونوع النموذج المستخدم

significance	Df	χ^2	Standardized Residual الباقي المعياري	Residual الباقي	Expected Frequencies التكرارات المتوقعة	Observed Frequencies التكرارات الملاحظة	INDICOTER المؤشر	الرتبة	توزيع القدرة
0.013	6	16.194	2.17	4.57	4.43	9	RMR	1	التواء سالب
			1.70	3.57	4.43	8	GFI	2	
			1.22	2.57	4.43	7	AGFI	3	
			-0.68	-1.43	4.43	3	RMSEA	4	
			-1.15	-2.43	4.43	2	TLI	5	
			-1.63	-3.43	4.43	1	CFI	6	
			-1.63	-3.43	4.43	1	IFI	7	
			0.00	0.00	31.00	31	الكلي		
0.082	11	18.000	1.79	4.00	5.00	9	AGFI	1	طبيعي
			1.79	4.00	5.00	9	GFI	2	
			1.79	4.00	5.00	9	RMR	3	
			1.34	3.00	5.00	8	RMSEA	4	
			0.00	0.00	5.00	5	TLI	5	
			-0.89	-2.00	5.00	3	CFI	6	
			-0.89	-2.00	5.00	3	IFI	7	
			-0.89	-2.00	5.00	3	HOELTER 0.01	8	
			-0.89	-2.00	5.00	3	HOELTER 0.05	9	
			-0.89	-2.00	5.00	3	PGFI	10	
			-0.89	-2.00	5.00	3	$\chi2$ /DF	11	
			-1.34	-3.00	5.00	2	PCFI	12	
			0.00	0.00	60.00	60	الكلي		
0.155	10	14.423	1.97	4.27	4.73	9	RMR	1	التواء موجب
			1.51	3.27	4.73	8	AGFI	2	
			1.51	3.27	4.73	8	GFI	3	
			0.59	1.27	4.73	6	RMSEA	4	
			0.13	0.27	4.73	5	TLI	5	
			-0.79	-1.73	4.73	3	CFI	6	
			-0.79	-1.73	4.73	3	IFI	7	
			-0.79	-1.73	4.73	3	HOELTER 0.01	8	
			-0.79	-1.73	4.73	3	$\chi2$ /DF	9	
			-1.25	-2.73	4.73	2	HOELTER 0.05	10	
			-1.25	-2.73	4.73	2	PGFI	11	
			0.00	0.00	52.00	52	الكلي		

بالرجوع للجدول 15 يلاحظ أن قيمة المؤشر χ^2 كانت دالة إحصائيا عند مستوى الدلالة (α = 0.05) وبدرجات حرية 6 لصالح المؤشر RMR، حينما كان شكل توزيع القدرة ملتو التواءً سالباً، حيث بلغت قيمة الباقي المعياري عندها (2.17) وهي أعلى من القيمة المتفق عليها. ولم تبدي المؤشرات الأخرى أي تطابق لشكل توزيع القدرة في حالة التوزيع الطبيعي. على الرغم أن المؤشرات AGFI، GFI، RMR، RMSEA حققت قيما موجبة مرتفعة في حساب الباقي المعياري حينما توزعت القدرة بشكل طبيعي، إلا أنها لم تحقق قيم التطابق المتفق عليها. وهذا يؤكد فاعلية المؤشر RMR في الكشف عن افتراض أحادية البعد في حالة الالتواء السالب.

أما في حالة شكل التوزيع الملتو التواءً موجباً كانت قيمة الباقي المعياري (1.97) لصالح المؤشر RMR ، وهذه القيمة تحقق التطابق عند مستوى الدلالة (α = 0.05). وتؤكد على فاعلية هذا المؤشر في الكشف عن افتراض أحادية البعد في حالة الالتواء الموجب. لذا يوصى بالمؤشر RMR كونه أكثر فاعلية في الكشف عن افتراض أحادية البعد في حالة تغيير شكل توزيع القدرة. فهذا المؤشر لا يتأثر كثيرا بتغيير شكل توزيع القدرة.

وعادة يوصى الباحثون هنكلز، ويرزما وجورس (Hinkle, Wiersma, & Jurs, 1988) بالمؤشرات ذات القيم الموجبة في حساب الباقي المعياري، كمؤشرات فاعلة في استخدامها في مثل تلك الحالات والتي لا تتأثر بتغيير شكل توزيع القدرة، والابتعاد ما أمكن عن المؤشرات ذات القيم السالبة في حساب الباقي المعياري، والتي تتأثر كثيرا بتغيير شكل توزيع القدرة.

ثالثا: النتائج المتعلقة بأثر شكل النموذج المستخدم على فاعلية المؤشرات المستخدمة في الكشف عن افتراض أحادية البعد.

للإجابة عن السؤال الثالث والذي ينص على: هل تختلف فاعلية مؤشرات الكشف عن تحقق افتراض أحادية البعد في النموذج البنائي المستخدم باختلاف النموذج (أحادي المعلمة، ثنائي المعلمة، ثلاثي المعلمة)؟ ولتحقيق ذلك تم حساب الاختبار الإحصائي χ^2 لحس المطابقة لكافة المؤشرات لدراسة أثر تغيير النموذج اللوجستي المستخدم وتحييد أثر طول الاختبار وشكل القدرة. وكانت النتائج موضحة بالجدول 16:

جدول 16: قيمة الإحصائي (χ^2) بتغيير نوع النموذج المستخدم وتحييد أثر طول الاختبار وشكل توزيع القدرة

significance	Df	χ^2	Standardized Residual الباقي المعياري	Residual الباقي	Expected Frequencies التكرارات المتوقعة	Observed Frequencies التكرارات الملاحظة	INDICOTER المؤشر	الرتبة	النموذج
0.365	11.0	11.986	1.27	3.08	5.92	9	AGFI	1	أحادي المعلم
			1.27	3.08	5.92	9	GFI	2	
			1.27	3.08	5.92	9	RMR	3	
			1.27	3.08	5.92	9	RMSEA	4	
			0.45	1.08	5.92	7	TLI	5	
			-0.38	-0.92	5.92	5	CFI	6	
			-0.38	-0.92	5.92	5	IFI	7	
			-0.79	-1.92	5.92	4	HOELTER 0.01	8	
			-0.79	-1.92	5.92	4	HOELTER 0.05	9	
			-0.79	-1.92	5.92	4	PGFI	10	
			-0.79	-1.92	5.92	4	$\chi2$ /DF	11	
			-1.61	-3.92	5.92	2	PCFI	12	
			0.00	0.00	71.00	71	الكلي		
0.216	3.0	4.462	0.98	2.50	6.50	9	RMR	1	ثنائي المعلم
			0.59	1.50	6.50	8	GFI	2	
			0.20	0.50	6.50	7	AGFI	3	
			-1.77	-4.50	6.50	2	RMSEA	4	
			0.00	0.00	26.00	26	الكلي		

significance	Df	χ^2	Standardized Residual الباقي المعياري	Residual الباقي	Expected Frequencies التكرارات المتوقعة	Observed Frequencies التكرارات الملاحظة	INDICOTER المؤشر	الرتبة	النموذج
			2.36	4.82	4.18	9	RMR	1	
			1.87	3.82	4.18	8	AGFI	2	
			1.87	3.82	4.18	8	GFI	3	
			0.89	1.82	4.18	6	RMSEA	4	
			0.40	0.82	4.18	5	TLI	5	
0.011	10.0	22.870	-1.07	-2.18	4.18	2	CFI	6	ثلاثي المعلم
			-1.07	-2.18	4.18	2	HOELTER 0.01	7	
			-1.07	-2.18	4.18	2	IFI	8	
			-1.07	-2.18	4.18	2	χ2 /DF	9	
			-1.56	-3.18	4.18	1	HOELTER 0.05	10	
			-1.56	-3.18	4.18	1	PGFI	11	
			0.00	0.00	46.00	46	الكلي		

بالرجوع للجدول 16 يلاحظ أن قيمة المؤشر χ^2 لحسن المطابقة كانت دالة إحصائيا عند مستوى الدلالة (α = 0.05) وبدرجات حرية (10) لصالح المؤشر RMR وفقا للنموذج ثلاثي المعلم، حيث كانت الدلالة الإحصائية (significance = 0.011)، وبلغت قيمة الباقي المعياري (2.36) وهي أعلى من القيمة المتفق عليها. وهذا يؤكد فاعلية المؤشر RMR في حالة النموذج ثلاثي المعلم. أما المؤشرات الأخرى لم تبدي أي تطابق لنوع النموذج المستخدم سواء في حالة أحادي البعد أو ثنائي البعد. حيث أن قيمة المؤشر χ^2 لحسن المطابقة كانت غير دالة إحصائيا عند مستوى الدلالة (α = 0.05) في الحالتين السابقتين.

بالرغم أن بعض المؤشرات (AGFI، GFI، RMR، RMSEA ، TLI) ، حققت قيما موجبة لبعض الحالات، إلا أنها لم تكن دالة إحصائيا لدى القيمة المتفق عليها. إنما يوصى باستخدامها في حالة اختلاف شكل النموذج المستخدم. كما يوصى بالابتعاد عن المؤشرات ذات القيم السالبة في حساب الباقي المعياري لتأثرها الكبير بتغيير شكل النموذج المستخدم.

كما تم حساب قيمة الاختبار الإحصائي (χ^2) للمؤشرات كافة، بغض النظر عن نوع المتغير المستقل؛ طول الاختبار (15، 30، 60) فقرة، شكل توزيع القدرة (ملتو التواء سالبا، طبيعي، ملتو التواء موجبا)، ونموذج استجابة الفقرة (أحادي المعلم،ثنائي المعلم، ثلاثي المعلم)، وكانت النتائج كما يلي:

جدول 17: قيمة الإحصائي (χ^2) للمؤشرات كافة وفقا لنماذج استجابة الفقرة باختلاف طول الاختبار وشكل توزيع القدرة

significance	Df	χ^2	Standardized Residual الباقي المعياري	Residual الباقي	Expected Frequencies التكرارات المتوقعة	Observed Frequencies التكرارات الملاحظة	INDICOTER المؤشر	الرتبة
0.000	11	74.091	4.37	15.08	11.92	27	RMR	1
			3.79	13.08	11.92	25	GFI	2
			3.50	12.08	11.92	24	AGFI	3
			1.47	5.08	11.92	17	RMSEA	4
			0.02	0.08	11.92	12	TLI	5
			-1.42	-4.92	11.92	7	CFI	6
			-1.42	-4.92	11.92	7	IFI	7
			-1.71	-5.92	11.92	6	HOELTER 0.01	8
			-1.71	-5.92	11.92	6	χ^2 /DF	9
			-2.004	-6.92	11.92	5	PGFI	10
			-2.004	-6.92	11.92	5	HOELTER 0.05	11
			-2.87	-9.92	11.92	2	PCFI	12
			0.00	0.00	143.00	143	الكلي	

بالرجوع للجدول 17 يلاحظ أن قيمة المؤشر χ^2 كانت دالة إحصائيا عند مستوى الدلالة ($\alpha = 0.05$) وبدرجات حرية (11) لصالح المؤشرات RMR، GFI، AGFI، وذلك بحساب الباقي المعياري، حيث كانت القيم أعلى من (2)، في حين أن المؤشرات PGFI، HOELTER 0.05، PCFI، كانت أقل من(2-). ويؤكد كل من هنكلز، ريزما و جورس (Hinkle, Wiersma, & Jurs, 1988) اعتماد قيمة الباقي المعياري كمعيار للحكم على فاعلية المؤشر.

وهذا يدل على أن المؤشرات ذات القيم الموجبة للباقي المعياري يوصى باستخدامها في الكشف عن افتراض أحادية البعد بغض النظر عن العوامل المؤثرة فيه. مما يعني أن هذه

المؤشرات (TLI ، RMSEA ، AGFI ،GFI ،RMR) أبدت فاعليتها في الكشف عن افتراض أحادية البعد. في حين أن المؤشرات ذات القيم السالبة للباقي المعياري لا يوصى باستخدامها.

ملخص لنتائج الدراسة.

توصلت الباحثة إلى ما يلي:

أولا: أن لطول الاختبار تأثير على فاعلية المؤشرات المستخدمة للتحقق من افتراض أحادية البعد، وأن طول الاختبار 30 فقرة قد حقق تطابقا بين النموذج المقترح والبيانات لدى المؤشرات الإحصائية AGFI، GFI، RMR، وأن 60 فقرة من طول الاختبار قد حقق تطابقا لدى مؤشرين فقط RMR، GFI. ولم يظهر مستوى طول الاختبار 15 فقرة أي تطابق مع جميع المؤشرات، وهذا يدل على أن مستوى 30 فقرة من طول الاختبار يتفق وافتراض أحادية البعد، بمعنى أن تغيير حجم الفقرات أثر على نتائج مؤشرات جودة المطابقة المستخدمة في الكشف عن افتراض أحادية البعد.

وهذا يعني أن المؤشرات التي ينصح باستخدامها للكشف عن افتراض أحادية البعد هي (GFI, RMR, AGFI)، وهي التي لا تتأثر كثيرا بتغيير طول الاختبار و المناسبة في الكشف عن افتراض أحادية البعد لدى تغيير طول الاختبار. وأن طول الاختبار المفضل في الكشف عن أحادية البعد هو 30 فقرة.

ثانيا: أن لشكل توزع القدرة أثر على فاعلية المؤشرات للكشف عن افتراض أحادية البعد، حيث وجدت الباحثة أن المؤشر RMR يصلح للكشف عن افتراض أحادية البعد حينما يكون شكل توزيع القدرة ملتو التواءً سالباً وملتو التواءً موجباً، وهذا يعني أن افتراض أحادية البعد يتحقق بصورة أكبر في حالة الالتواء لشكل توزيع القدرة، سيما أن شكل التوزيع

الطبيعي للقدرة لم يحقق افتراض أحادية البعد. ويوصى باستخدام المؤشر RMR للكشف عن افتراض أحادية البعد في حالة تغيير شكل توزيع القدرة.

ثالثا: للنماذج اللوجستية أثر على فاعلية المؤشرات المستخدمة في الكشف عن افتراض أحادية البعد، حيث أشارت النتائج لفاعلية المؤشر RMR في حالة البيانات المولدة من النموذج اللوجستي ثلاثي المعلم.

وحققت بعض المؤشرات (AGFI، GFI، RMR، RMSEA ، TLI) ، قيما موجبة في حساب الباقي المعياري في بعض الحالات، إلا أنها لم تكن دالة إحصائيا لدى القيمة المتفق عليها. إنما يوصى باستخدامها في حالة اختلاف شكل النموذج المستخدم. كما يوصى بالابتعاد عن المؤشرات ذات القيم السالبة في حساب الباقي المعياري لتأثرها الكبير بتغيير شكل النموذج المستخدم.

الفصل الخامس

مناقشة النتائج

تناول هذا الفصل مناقشة النتائج المتعلقة بأسئلة الدراسة، والتي تتعلق بالتحقق من فاعلية مؤشرات الكشف عن افتراض أحادية البعد وفق نماذج نظرية استجابة الفقرة، في ضوء تغير طول الاختبار وشكل توزيع القدرة.

مناقشة النتائج المتعلقة بمؤشرات جودة المطابقة.

- مناقشة النتائج المتعلقة بأثر طول الاختبار على فاعلية مؤشرات جودة المطابقة:

فيما يتعلق بأثر طول الاختبار على فاعلية المؤشرات المستخدمة للتحقق من افتراض أحادية البعد، حيث يتضمن هذا السؤال مقارنة أثر طول الاختبار (15، 30، 60) فقرة على فاعلية المؤشرات المستخدمة في الكشف عن أحادية البعد. فقد تمت دراسة كافة مؤشرات جودة المطابقة باستخدام التحليل العاملي التوكيدي وتم تبويبها بجداول خاصة كما هو موضح في فصل النتائج، وتم حساب الإحصائي χ^2 لحسن المطابقة لأثر تغيير طول الاختبار لكافة المؤشرات، فقد أشارت النتائج المتعلقة بقيم المؤشر χ^2 لحسن المطابقة أنها كانت دالة إحصائية عند مستوى الدلالة ($\alpha = 0.05$) بدرجات حرية 11 عند طول اختبار 30 فقرة وبمستوى احتمالية ($p = 0.006$)، لصالح المؤشرات AGFI، GFI، RMR، وقد حققت دلالة إحصائية ($p = 0.001$) وبدرجات حرية 11 عند مستوى 60 فقرة، ولصالح RMR، GFI. في حين أن قيمة المؤشر χ^2 لحسن المطابقة كانت غير دالة إحصائيا عند مستوى الدلالة ($\alpha = 0.05$) للمؤشرات جميعها عند مستوى طول الاختبار 15 فقرة.

وهذا يؤكد أن طول الاختبار 30 فقرة حقق تطابقا أكبر بين النموذج المقترح والبيانات لدى المؤشرات AGFI ، GFI، RMR، وأن 60 فقرة من طول الاختبار حقق تطابقا لدى

مؤشرين فقط RMR، GFI. ولم يظهر لطول الاختبار 15 فقـرة أي تطـابق مـع جميـع المؤشرات، وهذا يدل على أن 30 فقرة من طول الاختبار يتفق وافتراض أحادية البعـد، إذ أن تغيير حجم الفقرات أثر على نتائج مؤشرات جودة المطابقـة، بمعنـى آخـر أن هـذه المؤشرات لم تكن فاعلة في حالة عدد فقرات الاختبار القليل (15)، فـي حـين أن بعـض المؤشرات (RMR ، GFI) أبدت فاعليتها في الكشف عن افتراض أحادية البعد في حالة عدد الفقرات الكبير (60)، بينما حينما كان طول الاختبار (30) فقرة، كانت المؤشرات (AGFI، RMR ،GFI) أكثر فاعلية في الكشف عن افتراض أحادية البعد.

ويمكن القول أن القيم الموجبة في حساب البـاقي المعيـاري يمكـن أن يوصـى باستخدامها للكشف عن افتراض أحادية البعد لدى تغيير طول الاختبـار. ممـا يعنـي أن المؤشرات التي ينصح باستخدامها للكشف عن افتراض أحادية البعد هي (RMSEA, AGFI, GFI, RMR)، وذلك عندما يكون عدد فقرات الاختبار 30 فقرة، إذ لا تتأثر كثيرا بتغييـر طول الاختبار وهي المناسبة في الكشف عن افتراض أحادية البعـد لـدى تغييـر طـول الاختبار.

في حين أن المؤشرات الآتيـة (CFI، HOELTER 0.01، HOELTER 0.05، IFI، PGFI، PCFI) ذات القيم السالبة في حساب الباقي المعياري تتأثر بطول الاختبـار، ولـم تحقق نتائج للمطابقة بين النموذج المقترح والبيانات المولدة ولم تكن لها فاعلية في الكـشف عن افتراض أحادية البعد ولا ينصح باستخدامها. وهذا يتعـارض مـع دراسـة مكجيـل (McGill, 2009)، التي لم تكشف عن أثر طول الاختبار على دقة المؤشرات المستخدمة في الكشف عن افتراض أحادية البعد.

هذا يعني أن لطول الاختبار أثر على افتراض أحادية البعد حيث أكد كل من مـاك دونلاند ومولاك .(McDonald & Mulaik, 1979) أنه على الأقل لا بد من وجود 15 فقـرة

لقياس قدرة واحدة، في حين يشير كل من كارون ومايكل وكالن Karon , Michael & (Kallen,2009)، إلى أن حجم الفقرات الكبير (47) فقرة، كان أفضل من أعداد الفقرات القليلة (10) فقرة، في الكشف عن افتراض أحادية البعد. وأن بعض مؤشرات التحليل العاملي التوكيدي كانت أكثر فاعلية حيال الأعداد الكبيرة للفقرات في الكشف عن افتراض أحادية البعد. وقد أشار زانج (Zhang, 2008) بتأثر أساسيات افتراض أحادية البعد بطول الاختبار، فقد بين أن الاختبارات القصيرة أكثر عرضة لانتهاك افتراض أحادية البعد، وأن حجم الاختبار الكبير يقاوم انتهاك هذا الافتراض، وأن هذا الافتراض وثيق الصلة بتوازن أو تجانس المحتوى.

وترى الباحثة تبعاً للنتائج التي تم الحصول عليها أن حجم الفقرات المتوسط (30) فقرة يتفق مع افتراض أحادية البعد بصورة كبيرة ويزيد من فاعلية مؤشرات جودة المطابقة. وهذا يعود إلى أن كمية المعلومات التي تعطيها تلك الفقرات (30) كافية لتغطية متصل السمة. وأن زيادة حجم الفقرات لا يزيد من كمية معلومات الاختبار بل يبقيها ثابتة. إذ أن لكمية المعلومات صلة وثيقة بافتراض أحادية البعد. فقد أشار هامبلتون وسوامنثيان (Hambelton and Swaminathan, 1985) انه كلما زاد عدد الفقرات تزداد قيمة معلومات الاختبار، ويقل مقدار خطأ القياس. وهذا يزيد من فاعلية مؤشرات الكشف عن افتراض أحادية البعد ويعتبر مؤشراً على افتراض أحادية البعد.

ولم تتعرض أي من الدراسات السابقة لأثر طول الاختبار على افتراض أحادية البعد مما يجعل هذه الدراسة تتميز عن غيرها في تحقيق نتائج حول أثر طول الاختبار على فاعلية مؤشرات الكشف عن افتراض أحادية البعد.

- مناقشة النتائج المتعلقة بأثر شكل توزيع القدرة على فاعلية مؤشرات جودة المطابقة:

فيما يتعلق بأثر شكل توزيع القدرة على فاعلية المؤشرات المستخدمة للتحقق من افتراض أحادية البعد، والذي يتضمن مقارنة اثر توزيع شكل القدرة (التوزيع الطبيعي، ملتوٍ التواء موجب، ملتوٍ التواء سالب) على فاعلية المؤشرات المستخدمة في الكشف عن أحادية البعد.

فقد تبين من خلال النتائج أن قيمة المؤشر χ^2 قد كانت دالة إحصائيا عند مستوى الدلالة ($\alpha = 0.05$) وبدرجات حرية 6 لصالح المؤشر RMR حينما كان شكل توزيع القدرة ملتو التواءً سالباً، حيث بلغت قيمة الباقي المعياري عندها (2.17) وهي أعلى من القيمة المتفق عليها في الأدب النظري أكبر من 2+ وأصغر من 2-) (Hinkle, Wiersma, & Jurs, 1988).

أما بقية المؤشرات لم تبدي أي فاعلية لشكل توزيع القدرة في حالة التوزيع الطبيعي في الكشف عن افتراض أحادية البعد، على الرغم أن المؤشرات GFI، AGFI، RMR، RMSEA حققت قيما موجبة مرتفعة في حساب الباقي المعياري عندما توزعت القدرة بشكل طبيعي، إلا أنها لم تحقق قيم التطابق بين البيانات والنموذج المقترح تبعا لعلامة الباقي المعياري المشار إليها.

وقد كانت قيمة المؤشر χ^2 لحسن المطابقة (1.97)، عند شكل توزيع القدرة الملتو التواءً موجباً وهذه القيمة تحقق التطابق عند مستوى الدلالة ($\alpha = 0.05$). بمعنى أن الالتواء الموجب والالتواء السالب لشكل توزيع القدرة يتفق مع افتراض أحادية البعد، وأن شكل التوزيع الطبيعي للقدرة لم يحقق افتراض أحادية البعد. كما أن المؤشر RMR أظهر فاعليته في الكشف عن افتراض أحادية البعد، فهو لا يتأثر كثيرا بتغيير شكل توزيع القدرة. لذا يوصى بالمؤشر RMR للكشف عن افتراض أحادية البعد في حالة تغيير شكل توزيع القدرة.

وقد ذكر كارن ومايكل وكـــالن (Karon, Michael, & Kallen, 2008) أن طبيعـــة توزيع البيانات تؤثر على افتراض أحادية البعد، وأن التوزيع الملتـــوي الـــسالب لقـــدرات المفحوصين يحقق افتراض أحادية البعد. في حين لم تجد الباحثة أي اخـــتلاف فـــي تـــأثير الالتواء السالب عن الالتواء الموجب، فكليهما يتفق وافتراض أحادية البعد. كما يزيـــد مـــن فاعلية مؤشرات الكشف عن افتراض أحادية البعد كما هو الحال في مؤشر RMR.

ولتفسير ذلك تعزو الباحثة السبب إلى أن الالتواء السالب والالتواء الموجب يكـــون لمعلم الصعوبة، وأن انحياز الفقرات في اتجاه واحد دليل على الارتباطات المرتفعـــة فيمـــا بينها، وأن ترابطها معا يشكل اتساقاً داخلياً، وهذا يعطي مؤشراً إلى أن مجموعة الفقـــرات تقيس سمة أو قدرة واحدة، والارتباطات المرتفعة معا هي مؤشر علـــى الثبـــات المرتفـــع للاختبار. فقد أكد نانلي (Nunnally , 1979) على أن مؤشر الثبات دليـــل علـــى الاتـــساق الداخلي للفقرات وهو يشير إلى أحادية البعد.

وهذا يختلف ودراسة سوزان وزملائه.(Suzanne, et al., 2010) حيـــث أظهـــرت دراستهما تدني في قيم مؤشرات جودة المطابقـــة فـــي حالـــة الالتـــواء الـــسالب لقـــدرات المفحوصين. واختلفت مع دراسة كل من كيرسي وهسو ويو (Kirisci , Hsu & Yu, 1995) التي أظهرت نتائج مغايرة لاختلاف شكل توزيع قدرات المفحوصين، حيث أن شكل توزيع القدرة الملتو التواءً موجباً أعطى تقديرات ضعيفة في تقدير معلمة التخمـــين للفقـــرة، وأن البرامج الثلاثة (BILOG ، ASCAL ، MULTILOG) التي تم الاعتماد عليها؛ قدمت نتائج متشابهة عندما يكون شكل توزيع القدرة ملتو التواءً موجباً حول تقـــديرات معـــالم الفقـــرة للاختبار الأحادي. وأن التوزيع الطبيعي للقدرة، يقدم نتائج أفضل في تقدير معالم الفقرة.

كما اختلفت مع دراسة كل من كيرسي وهسو ويـــو (Kirisci , Hsu & Yu, 2001) التي بينت أن برامج التقدير أحادية البعد كانت أكثر حساسية لاختلاف شكل التوزيع وكانت مربعات الأخطاء لتقدير المعالم أقل ما يمكن في برنامج BILOG مقارنة بالبرامج الأخرى، وأظهر برنامج MULTILOG وبرنامج XCALIBRE تباينا قليلا في تقدير المعـــالم بـــسبب انتهاك افتراض أحادية البعد، وأن نتائجها كانت متناقضة.

وقد أشار كل من هيو وبلنتر (Hu & Bentler, 1999) الى أن المؤشر RMR لا يتأثر كثيرا بشكل توزيع القدرة، وأنه أكثر فاعلية في الكشف عن افتراض أحادية البعد، وعـــادة يوصى الباحثون (Hinkle, Wiersma, & Jurs, 1988) بالمؤشرات ذات القيم الموجبة للباقي المعياري في استخدامها في مثل تلك الحالات والتي لا تتأثر بتغيير شكل توزيـــع القـــدرة، والابتعاد ما أمكن عن المؤشرات ذات القيم السالبة، والتي تتأثر كثيرا بتغيير شكل توزيـــع القدرة.

ومما تجدر الإشارة إليه تفوق هذه الدراسة في تحقيق نتائج حول أثر شكل توزيـــع القدرة على افتراض أحادية البعد.

- مناقشة النتائج المتعلقة بأثر النموذج المستخدم وفقا لنماذج استجابة الفقرة على فاعلية مؤشرات جودة المطابقة:

وفيما يتعلق بأثر النموذج المستخدم على فاعلية المؤشرات المستخدمة للتحقق مـــن افتراض أحادية البعد، ويتضمن هذا السؤال أثر مقارنة النموذج (أحـــادي المعلـــم، ثنـــائي المعلم، ثلاثي المعلم) على فاعلية المؤشرات المستخدمة في الكشف عن أحادية البعد. حيث تم حساب الاختبار الإحصائي χ^2 لحسن المطابقة لكافة المؤشرات لدراسة أثر تغيير النموذج اللوجستي المستخدم.

أشارت النتائج أن قيمة المؤشر χ^2 لحسن المطابقة كانت دالة إحصائيا عند مستوى الدلالة ($\alpha = 0.05$) وبدرجات حرية (10) لصالح المؤشر RMR وفقا للنموذج ثلاثي المعلم، حيث كانت الدلالة الإحصائية (.011)، وبلغت قيمة الباقي المعياري (2.36) وهي أعلى من القيمة المتفق عليها (Hinkle, Wiersma, & Jurs, 1988)، وهذا يؤكد فاعلية المؤشر RMR في حالة البيانات المولدة من النموذج اللوجستي ثلاثي المعلم.

أما المؤشرات الأخرى لم تبدي أي تطابق لنوع النموذج المستخدم سواء في حالة النموذج أحادي المعلم أو ثنائي المعلم. حيث كانت قيمة المؤشر χ^2 لحسن المطابقة غير دالة إحصائيا عند مستوى الدلالة ($\alpha = 0.05$) في الحالتين السابقتين. على الرغم أن بعض المؤشرات (AGFI، GFI، RMR، RMSEA ، TLI) ، حققت قيما موجبة لبعض الحالات، إلا أنها لم تكن دالة إحصائيا لدى القيمة المتفق عليها. إنما يوصى باستخدامها في حالة اختلاف شكل النموذج المستخدم. كما يوصى بالابتعاد عن المؤشرات ذات القيم السالبة (CFI، HOELTER 0.01، HOELTER 0.05، IFI، PGFI، PCFI، χ^2 / df) لتأثرها الكبير بتغيير شكل النموذج المستخدم.

وتعزو الباحثة سبب تطابق البيانات المولدة من النموذج ثلاثي المعلم وافتراض أحادية البعد، هو أن هذا النموذج أكثر عمومية، ويشتمل على النموذجين أحادي المعلم وثنائي المعلم، فقد أشار هيولن ودراسغو وبارسونز (Hulin, Drasgow, & Parsons, 1983) إلى أن النموذج الثلاثي المعلم أكثر النماذج عمومية لأنه يسمح باختلاف معالم الفقرة الثلاثة، وبهذا يجب أن يؤخذ ذلك بعين الاعتبار عند مطابقة البيانات المستمدة من الاختبار لهذا النموذج.

واختلفت هذه النتيجة ودراسة ليو (Liu, 1992) التي أظهرت نتائجها أن البيانات المولدة من النموذج غير التعويضي تميل أكثر لأن تكون ثنائية البعد، وغير مطابقة للنموذج

أحادي البعد ثلاثي المعلم، وأن البيانات المولدة من النموذج التعويضي تميل لأن تكون أحادية البعد، وغير مطابقة للنموذج أحادي البعد الثلاثي المعلم، كما اختلفت ودراسة وانج ووانج (Wang & Wang, 2003) التي لم تظهر أي دلالة لأبعاد السمة (أحادي البعد، ثنائي البعد، ثلاثي البعد) على افتراض أحادية البعد. في حين أنها تتفق ودراسة كيرسي وهسو (Kirisci & Hsu, 1995) حيث أكدا أن بعدية الاختبار لها تأثير دال في تقدير معالم الفقرة والقدرة، حيث قدم الاختبار ثلاثي الأبعاد تقديرات أفضل في تقديره لمعلمة التخمين للفقرة.

ولدى دراسة الاختبار الإحصائي χ^2 لحسن المطابقة لكافة المؤشرات بغض النظر عن نوع المتغير المستقل؛ طول الاختبار (15، 30، 60) فقرة، شكل توزيع القدرة (ملتو التواء سالب، طبيعي، ملتو التواء موجب)، ونموذج استجابة الفقرة (أحادي المعلم،ثنائي المعلم، ثلاثي المعلم)، تبين أن قيمة المؤشر χ^2 كانت دالة إحصائيا عند مستوى الدلالة ($\alpha = 0.05$) وبدرجات حرية (11) لصالح المؤشرات RMR، GFI، AGFI، باعتماد قيمة الباقي المعياري كمعيار للحكم. حيث كانت القيم أعلى من (2)، في حين أن المؤشرات PGFI، HOELTER 0.05، PCFI، كانت أقل من 2- (Hinkle, Wiersma, & Jurs, 1988).

ويوصى باستخدام المؤشرات ذات القيم الموجبة بغض النظر عن العوامل المؤثرة في افتراض أحادية البعد، بمعنى آخر أن المؤشرات (RMR، GFI، AGFI، RMSEA، TLI)، تصلح كمؤشرات للكشف عن أحادية البعد، كونها أقل تأثر بتغيير حجم الاختبار وشكل توزيع القدرة ونوع النموذج المستخدم وكانت أكثر فاعلية في الكشف عن افتراض أحادية البعد. في حين أن المؤشرات ذات القيم السالبة لقيمة الباقي المعياري وهي: CFI، IFI، HOELTER 0.01، χ^2 /dF ، PGFI، HOELTER 0.05، PCFI، لا يوصى باستخدامها في الكشف عن افتراض أحادية البعد (Hinkle, Wiersma, & Jurs, 1988).

واختلفت مع دراسة كارون ومايكل وكـالن (Karon , Michael & Kallen,2009) التي تؤكد على فاعلية مؤشر المطابقة المعياري CFI، في حين بينت ان مؤشر جذر متوسط مربع الخطأ التقريبي RMSEA كان ضعيفا حيال الكشف عن أحادية البعد، ولم يكشف عن اثر حجم الفقرات على اثر تعدد الأبعاد. واتفقت في جزئية واحدة وهي فاعلية مؤشر توكر لويس TLI. كما اختلفـت ودراسـة كـل مـن هـامبلتون وروفنيـل (Hambleton & Rovinelli,1986)، بحيث لم يستطيعا الكشف عن افتراض أحادية البعد بنجاح.

في حين اتفقت ودراسة كيرسي وهسو (Kirisci & Hsu, 1995) حيث زود مؤشـر جذر متوسط مربع الخطأ التقريبي (RMSE) نتائج أفضل في حالة الاختبار أحادي البعد. في حين لم تزود دراسة كل من كرسي وهسو ويو (Kirisci , Hsu & Yu, 2001) أي معلومات عن فاعلية مؤشر الجذر التربيعي لمتوسط مربعات الأخطاء للفروق بين قيم المعالم المقدرة والمعالم الحقيقية، في الكشف عن افتراض أحادية البعد بصورة واضحة.

وتوصي الباحثة باسـتخدام المؤشـرات الآتيـة (RMR، GFI، AGFI، RMSEA، TLI)، وهي المؤشرات الأكثر فاعلية في الكشف عن افتراض أحادية البعد بغض النظر عن العوامل المؤثرة في افتراض أحادية البعد، بمعنى آخر أن تلك المؤشرات تصلح كمؤشرات للكشف عن أحادية البعد، وهي ذات أقل تأثر بحجم الاختبار وشكل توزيع القـدرة ونـوع النموذج المستخدم.

وتميل الباحثة إلى استخدام المؤشر RMR، للكشف عن افتراض أحادية البعد كون هذا المؤشر لا يتأثر كثيرا بتغيير طول الاختبار وشكل توزيع القدرة، كما انه أكثر فاعليـة لدى تغيير النموذج اللوجستي المستخدم ولا تتأثر بتغيير حجم المفحوصين، فـي حـين أن المؤشرات GFI، AGFI، RMSEA، تتأثر كثيرا بحجوم العينات وينصح بعدم اسـتخدامها كون الباحثة قد استخدمت حجم عينة كبير (1000) في دراستها (Hu & Bentler, 1999).

ولم تكشف أي من الدراسات السابقة حول أثر النموذج اللوجستي في التحقق من افتراض أحادية البعد، ولم تكشف عن فاعلية المؤشرات المستخدمة للكشف عن افتراض أحادية البعد في حالة تغيير النموذج اللوجستي، وهذا ما تميزت به هذه الدراسة.

التوصيات.

من خلال النتائج التي توصلت إليها الدراسة الحالية يمكن للباحثة اقتراح التوصيات التالية:

1- إجراء المزيد من الدراسات حول التحليل العاملي التوكيدي للكشف عن افتراضات نظرية استجابة الفقرة وأهميتها في تقدير معالم الفقرة والمفحوصين.

2- إجراء دراسات حول أثر عدد المفحوصين في الكشف عن افتراضات نظرية استجابة الفقرة.

3- الاهتمام بالدراسات المتعلقة بالتحليل العاملي التوكيدي بصورة أكبر والمنبثقة عن نموذج المعادلة البنائية وأثرها في التحقق من افتراضات نظرية استجابة الفقرة، إضافة لأهميتها في إعطاء تقديرات أفضل لمعالم الفقرة وتقديرات قدرة المفحوصين، وأهميتها في إعداد بنوك الأسئلة ومعايرتها. ويوصى باستخدام مؤشرات التحليل العاملي التوكيدي أثناء إنشاء وبناء بنوك الأسئلة، للتحقق من افتراضات نظرية استجابة الفقرة. وهذا الأمر خاص بمعدي ومطوري الاختبارات في مراكز الاختبارات الكبيرة.

4- يوصى بإجراء المزيد من الدراسات المتعلقة بالتحقق من افتراض أحادية البعد باسلوب التحليل العاملي التوكيدي لدى النماذج اللوجستية الأحادية البعد والمتعددة التدريج.

5- إجراء المزيد من الدراسات المتعلقة بالتحقق من افتراض أحادية البعد باسلوب التحليل العاملي التوكيدي على بيانات حقيقية، إذ اقتصرت الدراسة الحالية على البيانات المولدة.

المراجع العربية:

علام، صلاح الدين. (2005) **نماذج الاستجابة للمفردة الاختبارية أحادية البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي**، دار الفكر العربي.

الشافعي، محمد منصور. (2008). **تأثير انتهاك افتراض أحادية البعد واستقلالية المحل في تدريج بنك الأسئلة ودقة معادلة درجات الاختبارات البنكية المسحوبة**، ورقة عمل، الندوة الإقليمية لعلم النفس، كلية التربية، جامعة الملك سعود، مطبوعات الندوة.

المراجع الأجنبية:

Ackerman, A. (1994). Using Multidimensional item response theory to under stand what items and tests are measuring. *Applied measurement in Education, 7(4), 255 – 278.*

Ajzen, I., & Madden, T. J. (1986). Prediction of goal-directed behavior: Attitudes, intentions, and perceived behavioral control. *Journal of Experimental Social Psychology*, *22*, 453-474.

Allen, M.J, and Yen, W.M. (1979). *Introduction to Measurement Theory***.** California: Cole publishing company.

Anastasi, A. & Urbina, S. (1997). *Psychological Testing. (7 ed.),* New York: Printice Hall.

Anglim, J. (2007).*Structural Equation Modelling*. (325-711) Research Methods. psych. unimelb.edu.au/jkanglim.

Baker, F. (2001). The basics of item response theory. *ERIC Clearinghouse on Assessment and Evaluation*, University of Maryland, College Park, MD.

Bentler, P. M. and C. P. Chou (1987). Practical issues in structural modeling. *Sociological Methods and Research*. 16(1): 78-117.

Bentler, P.M., & Bonett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88,* 588-606.

Bollen, K.A. (1986). *Sample size and Bentler and Bonett's nonnormed fit index*. Psychometrika, 51, 375–377.

Bollen, K. A. (1989). *Structural Equations with Latent Variables*. New York: Wiley.

Bollen, K. A. (1990). Overall fit in covariance structure models: Two types of sample size effects. *Psychological Bulletin* 107(2): 256-259.

Bollen, K.A. & Long, J.S. [Eds.] (1993). *Testing structural equation models*. Newbury Park, CA: Sage.

Byrne, Barbara (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS*. Hillsdale, NJ: Lawrence Erlbaum. Mainly covers LISREL 8, but also PRELIS 2, and SIMPLIS. Some emphasis on but not limited to psychology.

Carmines*, E. G. &* Zeller, R. A. (1979) *Reliability and Validity Assessment*. Beverly. Hills, CA: Sage. Chubb, John E. Multiple Indicators .

Crocker, L. & Algina, J. (1986) *INTRODUCTION TO CLASSICAL AND MODERN TEST THEORY* .New York.

Daniel ,M,Bolt , Allans Cohen , Jams, A.(2002). Item Parameter Estimation Under Conditions Of Test Speediness: Application Of Mixture Rasch Model With Ordinal Constraints *, Journal Of Educational Measurement*, Vol(39),Issue(4),December ,.pp.(331-348).

Douglas,m.l., (2006). The Presence And Impact Of Local Item Dependence On Objective Structured Clinical Examinations Scores And The Potential Use Of The Polytomous,Many-facet Rasch Model*, Applied Psychological Measurement* , Vol (146), N(25).

Duncan, O. D.(1975). *Introduction to Structural Equation Models*. New York: Academic Press.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists.* Mahwah, NJ: Erlbaum.

Fox, J. (1984). Linear Statistical Models and Related Methods: *With Applications to Social Research*. NewYork: Wiley.

Glorfeld, L.W. (1995). An improvement on Horn's parallel analysis methodology for selecting the correct number of factors to retain. *Educational and Psychological Measurement. 55,* 377-393.

Greene, W. H. (1993). *Econometric Analysis*, Second Edition. New York: Macmillan.

Gustafsson, J. (1980). Testing and obtaining fit of data to the Rasch model. *British Journal of Mathematical and Statistical Psychology, 33, 205 – 233.*

James, L.R., Mulaik, S.A. & Brett, J.M. (1982). *Causal analysis: Assumptions, models and data. Beverly Hills: Sage.*

Judge, G. G., W. E. Griffiths, R. C. Hill, H. Lütkepohl & T.-C. Lee.(1985). *The Theory and Practice of Econometrics*, Second Edition. New York: Wiley.

Hambleton, R., & Rovinelli. R. (1986). Assessing the dimensionality of a set of test item. *Apaper presented at the annual meeting of AERA, Montreal*.

Hambleton, R.K., & Jones, R.W. (1993). Comparison of Classical Test Theory and Item Response Theory and their Applications to test development. *Educational Measurement***.***Issues and Practice***,** 38-47.

Hambleton, R. K., & Swaminathan, H. (1985) *Item Response Theory. Principles and Applications* . kluwer. Nijhoff publishing.

Hambleton, R. K., Swaminathan, H. & Rogers H. J. (1991) *Fundamentals of Item Response Theory*. New York, NY: Sage Publications.

Hattie, J. (1985). Methodology review: Assessing unidimensionality of tests and items. *Applied Psychological Measurement*, *9*, 139–164.

Hinkle, D, E. Wiersma, W. & Jurs, S. G. (1988). *Citation of applied statistics for the behavioural sciences*, 2nd edition , Houghton Mifflin Company

Hulin, C. L., Drasgow, F. & Parsons, C.K. (1983). Item Response Theory: *Application to Psychological Measurement*. Illinois: Dow Jones _ Irwin.

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 4*, 424-453.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55.

Hoelter, J.W. (1983). The analysis of covariance structures: Goodness-of-fit indices. *Sociological Methods and Research*, 11, 325–344.

Hooper, D., Coughlan, J. and Mullen, M. R.(2008 *) Structural Equation Modelling*: Guidelines for Determining Model Fit. *The*

Karon F. Cook Michael A. Kallen.(2009). *Having a fit: impact of number of items and distribution of data on traditional criteria for assessing IRT's unidimensionality assumption*. Qual Life Res (2009) 18:447–460 DOI 10.1007/s11136-009-9464-4.

Kenny, D. A., & McCoach, D. B. (2003). Effect of the number of variables on measures of fit in structural equation modeling. *Structural Equation Modeling, 10*, 333-3511.

Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2011). *The performance of RMSEA in models with small degrees of freedom.* Unpublished paper, University of Connecticut.

Kenny, D. (2012). *Measuring Model Fit*. published paper, http://davidakenny.net/cm/fit.htm

Kline R. B. (1998). Software programs for structural equation modeling: AMOS, EQS, and LISREL. *Journal of Psychoeducational Assessment* (16): 343-364.

Kirisci , L. & Hsu, T. (1995*). The Robustness of BILOG to violations the Assumptions of Unidimensionality of test items and normality of ability distribution*. Paper presented at the annual meeting of the national council on Measurement in education, San Francisco, CA, April, 19 * 21, 1995 (ERIC Document Reproducation Service No. ED 384646).

Kirisci , L. Hsu, T & Yu, Life. (2001). Robustness of Item Parameter Estimation Programs to Assumptions of Unidimensionality and Normality. *Applied Psychological Measurement; Jun2001, Vol. 25 (2), 146 – 162.*

Knol, D. L., & Berger, M. P. F. (1991). Empirical comparison between factor analysis and multidimensional item response models. *Multivariate Behavioral Research, 26*, 457-477.

Lee, Guemin. (2000). Estimating conditional standard error of measurement for tests composed of trestles. *Applied Measurement In Education*, 13(2), 161 – 180.

Lee, S. & Terry, R. (2005). *MDIRT - FIT: SAS Macros for fitting multidimensional item response*. Presented at SUGI 31th conference in University of Oklahoma, Norman, Ok.

Liu, X. (1992). *The Dimensionality of test data generated by compensatory and non- compensatory two dimensional IRT models and its effects on model data fit*. Paper presented at the annual meeting of the Canadian Society for the study of Education (Charlottetown, Prince Edward Island, June, 1992).

Lord, F . (1980) *Applications of Item Response Theory to Practical Testing Proplems*. Hillsdale, N.J. : Lawrence Erlbaum.

Linacre, J.M. (2008). Winsteps Rasch measurement (Version 3.63.2). Chicago, IL: MESA Press. Masters. N.G. (1982). *A Rasch Model For Partial Credit Scoring*. Psychometrika, 47, 149-174 .

McDonald, R. P., & Muliak, S. (1979). Determinancy of common factors: A non-technical review. *Psychological Bulletin, 86*, 297-306.

McGill. M. T (2009**).** *An Investigation of Unidimensional Testing Procedures under Latent Trait Theory using Principal Component Analysis***.**

Mulaik, S.A., James, L.R., Van Alstine, J., Bennett, N., Lind, S. & Stilwell, C.D. (1989). Evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin*, 105, 430–445.

Nunnally, J. C. (1979).*Introduction to Psychological Measurement*. New York : McGraw – Hill.

Rasch, G. (1966). An item analysis which takes indivdual differences into account. *British Journal of Mathematical and Statistical Psychology, 19, 49 – 57.*

Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor tests: Results and implications. *Journal of Educational Statistics, 4,* 207–230.

Reckase, M. D. (1985). The difficulty of items that measure more than one ability. *Applied Psychological Measurement*, *9*, 401–412.

Reckase, M. D., & McKinley, R. L. (1991). The discriminating power of items that measure more than one dimension. *Applied Psychological Measurement, 15,* 361–373.

Reese, l.m. & Pashley.p.j (2007). *impact of local item dependence on true score design* ,LSAC, Research Report Series.

Reis, R. (1986). A structural equation analysis of Weiner's attribution-affect model of helping behavior. *Journal of Personality and Social Psychology, 50*, 1123-1133.

Reise, S.P., & Waller, N.G, (2003).How many IRT Parameters Does It Take To Model Psychopathology Items?, *Psychological Methods*,Vol (8), pp 164–184.

Stevens, J.P. (2002). *Applied multivariate statistics for the social sciences.* (4th ed), Lawrence Erlbaum: Mahwah, NJ.

Stout, W.R. (1987). A nonparametric approach for assessing latent trait unidimensionality. *Psychometrika, 52,* 589-617.

Suzanne L. Slocum-Gori, Bruno D. & Zumbo (2010). *Assessing the Unidimensionality of Psychological Scales:Using Multiple Criteria from Factor Analysis*. Soc Indic, 102:443–461.

Tanaka, J. S. (1987). How big is big enough? Sample size and goodness of fit in structural equation models with latent variables. *Child Development, 58,* 134-146.

Timothy, A. (2006). *Confirmatory Factor Analysis For Applied Research*. New York, NY, 10012.

Tinsley, H. E. & Dawis, R. V. (1974). An investigation of the rasch simple logistic model: Sample free item and test calibration. *Educational and Psychological Measurement, 11, 163 – 178.*

Walker, C. & Betetras, S. (2000). *Using Multidimensional versus Unidimensional ability estimates to determine student proficiency in mathematics*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA, April24 – 28, 2000.

Wang, Y. (2005). Robustness Of Unidimensional IRT Calibration In Presence Of Essential Dimensionalty, *ERIC*,-No:ed371019.

Wang, S.& Wang, N. (2003). *The Effects of Multidimensional Ploytomous Response Data on Unidimensional Many Fact Rasch Model Parameters Estimates*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL, April, 21 – 25, 2003.

Warm, T. (1979). *Primer of item Response Theory*. Oklahoma City: U. S. Coast Guard Institute.

Weiss, D. J. (1983). *New horizons in testing.* New York: Academic Press.

ABSTRACT

Alhawari, Arwa Issa Abed. (2013). Verifying the Effectiveness of Unidimensionality Assumption Indicators According to Item Response Theory Models (IRT), in light of test length and Ability distribution. Ph. D. Dissertation, Yarmouk University. (Supervisor: Dr. Nedal al-Shraifin).

This study aimed to verify the effectiveness of Unidimensionality assumption Indicators According to Item Response Theory Models (IRT), in light of test length and Ability distribution. By using confirmatory analysis. To achieve this aim, the concept of Unidimensionality assumption was defined, and its importance, and the consequences resulting from its violation, were illustrated. The required conditions for Unidimensionality were identified, and the Unidimensionality Indicators were explained. In addition, the Item Response Theory Models was defined, with their assumption, and the required conditions of confirmatory analysis. were identified too. Furthermore, explaining the most important indicators that resulted from the corresponding fit indicators, and defining the most important characteristics of such indicators. chi square test (χ^2) was also employed in the detection of the effectiveness of Unidimensionality Assumption.

The Structural Equation Model Fit Indices were presented, such as the ratio between the value of χ 2, degrees of freedom df, and absolute Fit Index. And incremental Fit Indexes, . In addition the Hoelter's (1983)`critical N' for a 0.05. according to criteria point agreed upon in the literature review.

The results of the confirmatory analysis indicated the importance of Unidimensionality assumption, and the effectiveness of the indicators utilized in the detection of Unidimensionality assumption. Among the procedures of this study was generating data, and generating an items were also generating the examinee's

ability, and generating data rate (1000) were examined for all variables; test length (15, 30, 60) item, and Ability distribution (normal, positive and negative skewed), according to Item Response Theory Models (one-parameter, two-parameter, three-parameter) and analysis of data generated Depending on the software AMOS 21. Also used the chi square test (χ 2) for detecting the effectiveness of matching fit indicators.

The results showed that the length of the test affected the Effectiveness of Unidimensionality assumption Indicators, and the test consisted from 30 items in order to achieve a good fitting between the proposed model and the data generated, especially the statistical indicators AGFI, GFI, RMR. Also, the Ability distribution impact on the effectiveness of the positive and negative skewed achieve Unidimensionality assumption Indicators, and the logistics models impact on the effectiveness of the indicators used, where the results showed the effectiveness of the RMR index in the case of the data generated from the logistic three-parameter model.

In light of the results researcher recommends using (RMR, GFI, AGFI), Indicator Fit Index followed the conformity analysis in the detection of the Unidimensionality assumption regardless of the factors influences, which were less affected by the length of the test and shape of the Ability distribution and the type of logistics models used.

Keywords: Unidimensionality assumption, conformity analysis, Structural Equation Model, Fit Indices, Item Response Theory Models, software AMOS 21.